The Dissertation Committee for Andrew David Grotzinger Certifies that this is the approved version of the following Dissertation:

GENOMIC STRUCTURAL EQUATION MODELING: A NOVEL METHOD USED TO UNDERSTAND THE SHARED GENETIC BASIS OF MENTAL HEALTH DISORDERS

Committee:

Elliot M. Tucker-Drob, Co-Supervisor

Kathryn Paige Harden, Co-Supervisor

Michel G. Nivard

Christopher G. Beevers

James P. Curley

GENOMIC STRUCTURAL EQUATION MODELING: A NOVEL METHOD USED TO UNDERSTAND THE SHARED GENETIC BASIS OF MENTAL HEALTH DISORDERS

by

Andrew David Grotzinger

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2021

Abstract

GENOMIC STRUCTURAL EQUATION MODELING: A NOVEL METHOD USED TO UNDERSTAND THE SHARED GENETIC BASIS OF MENTAL HEALTH DISORDERS

Andrew David Grotzinger, Ph.D.

The University of Texas at Austin, 2021

Co-supervisors: Elliot M. Tucker-Drob, Kathryn Paige Harden

Methods for using genome-wide association studies (GWAS) to estimate genetic correlations between pairwise combinations of traits have produced "atlases" of genetic architecture. Genetic atlases reveal pervasive pleiotropy, and genome-wide significant loci are often shared across a wide variety of phenotypes. The present dissertation consists of three main sections, with the overarching goal of developing and applying multivariate methods that allow greater understanding of psychiatric, genetic atlases. First, Genomic Structural Equation Modeling (Genomic SEM) is tested and validated. Using formal methods for modeling covariance structure, Genomic SEM synthesizes genetic correlations and SNP-heritabilities inferred from GWAS summary statistics of individual traits from samples with varying and unknown degrees of overlap. Second, Stratified Genomic SEM is developed as an extension that allows for modeling genetic covariance stratified across biologically meaningful classes of genes. These stratifications can include classes of genes expressed during certain developmental epochs (e.g., genes expressed prenatally), tissues (e.g., genes expressed in the hippocampus), and cell types (e.g., pyramidal neurons). Third, Genomic SEM and Stratified Genomic SEM are applied to provide a greater understanding of shared genetic risk across 11 major classes of

psychiatric traits. These results provide insight into genetic risk sharing across the diagnostic spectrum at the genome-wide level, for classes of genes, and for individual genetic loci.

Table of Contents

List of Tables

List of Figures

General Introduction

Approximately half of individuals with a mental disorder concurrently meet criteria for another disorder (Kessler, Chiu, Demler, & Walters, 2005), and ~41% of individuals will meet criteria for four or more disorders in their lifetime (Caspi & Moffitt, 2018). In addition, offspring of parents with mental illness are at higher risk for developing *any* mental disorder, and not just the specific parental disorder (Dean et al., 2018; Martel et al., 2017; McLaughlin et al., 2012). Bivariate genomic methods (e.g., GCTA & LD-score Regression) consistently reveal high levels of genetic covariance across a diverse range of psychiatric disorders (Bulik-Sullivan et al., 2015; Selzam, Coleman, Caspi, Moffitt, & Plomin, 2018). Thus, a substantive reason for non-specific familial risk, and pervasive psychiatric comorbidity more generally, appears to be widespread statistical pleiotropy.

In addition to phenotypic patterns of comorbidity and evidence for genetic overlap, the extant literature points to a number of overlapping biological mechanisms across psychiatric traits. With respect to neurotransmitters, dysregulations in the glutamatergic system are widely cited for attention-deficit/hyperactivity disorder (ADHD; e.g., Elia et al., 2011), bipolar disorder (e.g., Forstner et al., 2015), and schizophrenia (e.g., Hu et al., 2015). GABAergic abnormalities have been identified across an even wider range of disorders, including: schizophrenia (Guillozet-Bongaarts et al., 2014), bipolar disorder (Fatemi et al., 2013), major depressive disorder (MDD; Gao et al., 2013), ADHD (Fatemi et al., 2013), anxiety disorders (Mohler, 2012), and autism spectrum disorder (Fatemi et al., 2010). Structural abnormalities are also commonly reported across all disorders, but grey matter loss specific to the hippocampus may play a role in schizophrenia (Crossley et al., 2009), bipolar disorder (Hibar et al., 2017), and

MDD (Schmaal et al., 2016). Thus, genetic and, more generally, biologically informed research clearly points to a need for multivariate methods that can be used to more systematically investigate overlapping genetic pathways of psychiatric risk.

The current dissertation consists of three main parts. First, I develop and validate a novel method, Genomic Structural Equation Modeling (Genomic SEM; Grotzinger et al., 2019), that allows genetic covariance matrices produced from GWAS summary statistics to be formally modeled. In Chapter 2, I validate Stratified Genomic SEM, an extension of the Genomic SEM framework that allows genetic covariance to be modeled within biologically meaningful categories, allowing us to examine genetic enrichment for overarching factors of psychiatric liability. These biological categories are constructed based on collateral information from gene expression methods (e.g., RNA-Seq), and include classes of genes that map onto the overlapping biological mechanisms noted in the prior paragraph (e.g., genes expressed in the hippocampus). While independent studies that include individual diagnoses repeatedly point to the convergence of biological risk factors, Stratified Genomic SEM allows this risk sharing to best tested for the first time across an encompassing range of the diagnostic spectrum. In Chapter 3, I pair Genomic SEM with the most up-to-date summary statistics across 11 classes of major psychiatric disorders (average total sample size = 154,923 participants), to systematically evaluate competing structural models. I subsequently apply Stratified Genomic SEM to identify genetic enrichment of psychiatric factors within biologically meaningful categories of genes. Finally, I utilize multivariate GWAS within Genomic SEM to identify risk conferring loci across the psychiatric factors. Collectively, these results leverage genetic correlations, biologically relevant

pathways, and individual genetic variants, to offer critical and comprehensive insight into the nature of pervasive pleiotropy across psychiatric traits.

Chapter 1: Validation of Genomic SEM[1]

INTRODUCTION

Genome-wide association studies (GWASs) are rapidly identifying loci affecting multiple social, behavioral, and psychiatric phenotypes (Lee et al., 2013; Bush, Oetjens, & Crawford, 2016). Moreover, using cross-trait versions of methods such as genomic-relatedness-based restricted maximum-likelihood (GREML; Yang, Lee, Goddard, & Visscher, 2011) and LD-score regression (LDSC; Bulik-Sullivan et al., 2015) researchers have identified genetic correlations between diverse traits, such as age of first birth and risk of smoking (Barban et al., 2016), insomnia and psychiatric traits (e.g., schizophrenia; Jansen et al., 2018), and major depressive disorder and number of children (Wray et al., 2018). In fact, widespread pleiotropy appears to be the rule rather than the exception across phenotypes. Although these findings are currently suggestive of constellations of phenotypes affected by shared sources of genetic liability, existing methods do not permit the causes of the observed genetic correlations to be investigated systematically.

As part of this dissertation, I have worked actively with my supervisors to create and validate Genomic SEM, a novel method for modeling the multivariate genetic architecture of constellations of traits and incorporating genetic covariance structure into multivariate GWAS discovery. Genomic SEM is a flexible framework for formally modeling the genetic covariance

[1] Results reported in the first chapter appear in the published work (Grotzinger et al., 2019). My contributions to this work included the conceptualization, analyses, and write-up of the results.
Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., ... & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature human behaviour*, *3*(5), 513-525.

structure of complex traits using GWAS summary statistics from samples of varying and potentially unknown degrees of overlap, in contrast to existing methods that model phenotypic covariance structure (Verhulst, Maes, & Neale, 2017), with very specific applications (Beaumont et al., 2018), using raw data. Moreover, Genomic SEM allows for the specification and comparison of a range of different hypothesized multivariate genetic architectures, which improves upon existing approaches for combining information across genetically correlated traits to aid in discovery (Turley et al., 2018).

Genomic SEM is a Two-Stage Structural Equation Modeling approach (Cheung, 2015; Savalei & Bentler, 2009; Yuan & Bentler, 2000). In Stage 1, the empirical genetic covariance matrix and its associated sampling covariance matrix are estimated. The diagonal elements of the sampling covariance matrix are squared standard errors (*SE*s). The off-diagonal elements index the extent to which sampling errors of the estimates are associated, as may be the case when there is sample overlap across GWAS. In principle, these matrices may be obtained using a variety of methods for estimating SNP heritability, their genetic covariance, and their *SE*s. Here we use a novel, multivariate version of LDSC that accounts for potentially unknown degrees of sample overlap by populating the off-diagonal elements of the sampling covariance matrix. The same strengths, as well as assumptions and limitations, that are known to apply to LDSC (de Vlaming, Johannesson, Magnusson, Ikram, & Visscher, 2017; Lee & Chow, 2017) apply to its extension used here and to Genomic SEM. In Stage 2, a SEM is estimated by minimizing the discrepancy between the model-implied genetic covariance matrix and the empirical covariance matrix obtained in the previous stage.

Stage 1 Estimation

In Stage 1, the empirical genetic covariance matrix (S_{LDSC}) and its associated sampling covariance matrix ($V_{S_{LDSC}}$) are estimated using our multivariable extension of LDSC. S_{LDSC} is a $k \times k$ symmetric matrix with SNP heritabilities on the diagonal and genetic covariances ($\sigma_{gi,gj}$) between phenotypes *i* and *j* off the diagonal. The genetic covariance between phenotypes *i* and *j* can be computed as the genetic correlation scaled relative to the total genetic variance of each of the two contributing phenotypes (themselves scaled to unit variances), $\sigma_{gi,gj} = r_{gi,gj} \cdot \sqrt{h_i^2 \cdot h_j^2}$. Thus, the genetic covariance matrix of order *k* has $k^* = k\left(k+1\right)/2$ nonredundant elements. It can be written as:

$$S_{LDSC} = \begin{bmatrix} h_1^2 & & & \\ \sigma_{g1,g2} & h_2^2 & & \\ \vdots & & \ddots & \\ \sigma_{g1,gk} & \sigma_{g2,gk} & \cdots & h_k^2 \end{bmatrix}$$

To produce unbiased *SE* estimates and test statistics, we require the asymptotic sampling covariance matrix, $V_{S_{LDSC}}$, of the LDSC estimates that is composed of all nonredundant elements in the S_{LDSC} matrix. Thus, it is a symmetric matrix of order *k**, with $k^*\left(k^*+1\right)/2$ nonredundant elements. The diagonal elements of $V_{S_{LDSC}}$ are sampling variances, that is, squared *SE*s of the elements in S_{LDSC}. The off-diagonal elements of $V_{S_{LDSC}}$ are sampling covariances that indicate the extent to which the sampling distributions of the variance and covariance estimates in S_{LDSC} covary with one another, as would be expected when there is overlap among the samples from which the terms are estimated. This $V_{S_{LDSC}}$ matrix can be written as:

$$V_{S_{LDSC}} = \begin{bmatrix} SE(h_1^2)^2 & & & & & & \\ \text{cov}(h_1^2, \sigma_{g1,g2}) & SE(\sigma_{g1,g2})^2 & & & & & \\ \vdots & \vdots & \ddots & & & & \\ \text{cov}(h_1^2, \sigma_{g1,gk}) & \text{cov}(\sigma_{g1,g2}, \sigma_{g1,gk}) & SE(\sigma_{g1,gk})^2 & & & & \\ \text{cov}(h_1^2, h_j^2) & \text{cov}(\sigma_{g1,g2}, h_j^2) & \text{cov}(\sigma_{g1,gk}, h_j^2) & \ddots & & & \\ & & & SE(h_j^2)^2 & \ddots & & \\ \text{cov}(h_1^2, \sigma_{gj,gk}) & \text{cov}(\sigma_{g1,g2}, \sigma_{gj,gk}) & \text{cov}(\sigma_{g1,gk}, \sigma_{gj,gk}) & \text{cov}(h_j^2, \sigma_{gj,gk}) & SE(\sigma_{gj,gk})^2 & & \\ \text{cov}(h_1^2, h_k^2) & \text{cov}(\sigma_{g1,g2}, h_k^2) & \text{cov}(\sigma_{g1,gk}, h_k^2) & \text{cov}(h_j^2, h_k^2) & \text{cov}(\sigma_{gj,gk}, h_k^2) & SE(h_k^2)^2 \end{bmatrix}$$

The diagonal elements of $V_{S_{LDSC}}$ are estimated using a jackknife resampling procedure in the bivariate version of LDSC that is currently available by its original developers (Bulik-Sullivan et al., 2015). The LDSC function introduced in the software package expands the jackknife procedure to the multivariable context in order to produce sampling covariances (which index dependencies among estimation errors) among the elements of S_{LDSC}, needed to populate the off-diagonal elements of $V_{S_{LDSC}}$.

Incorporation of Individual SNP Effects

Several steps are needed to incorporate individual SNP effects into Genomic SEM. The first step requires that the inputted genetic covariance matrix be expanded to include covariances between the SNP and each of the phenotypes, g_1 through g_k, by appending a vector of SNP-phenotype covariances (S_{SNP}) to S_{LDSC}:

$$S_{Full} = \begin{bmatrix} \sigma^2_{SNP} & & & & & \\ \sigma_{SNP,g1} & h^2_1 & & & & \\ \sigma_{SNP,g2} & \sigma_{g1,g2} & h^2_2 & & & \\ \sigma_{SNP,g3} & \sigma_{g1,g3} & \sigma_{g2,g3} & h^2_3 & & \\ \vdots & \vdots & & & \ddots & \\ \sigma_{SNP,gk} & \sigma_{g1,gk} & \sigma_{g2,gk} & \sigma_{g3,gk} & \cdots & h^2_k \end{bmatrix}$$

The sampling covariance matrix, V_{SFull}, associated with this expanded S_{Full} covariance matrix includes a number of components. One block of this V_{SFull} matrix, $V_{S_{LDSC}}$, contains the sampling variances and sampling covariances of the latent genetic variances (SNP heritabilities) and genetic covariances, which are obtained from the multivariable LDSC approach introduced above. A second block of the V_{SFull} matrix, V_{SSNP}, is composed of the sampling covariance matrix of the SNP effects on the phenotypes. The SNP variance (derived from reference panel data) is treated as fixed, and its sampling variance and sampling covariance with all other terms are fixed to 0 (or to a very small value to facilitate computational tractability). The sampling covariances of the SNP-genotype covariances with one another are obtained using cross-trait LDSC intercepts (which represent sampling correlations weighted by sample overlap) after being rescaled relative to the sampling variances of the respective SNP-genotype covariances (Baselmans et al., 2017; Turley et al., 2018). A final block of the V_{SFull} matrix represents the sampling covariance of the SNP-genotype covariances with the genetic variances and genetic covariances. These are fixed to 0, as sampling variation of the SNP-genotype covariance is expected to be independent of the test statistics of all LD blocks except the one it occupies. Because the sampling variance of the heritabilities and genetic correlations derive from sampling variability in the test statistics within all of the LD blocks, their sampling covariances with a

single SNP effect is expected to approach 0. In sum, the V_{SFull} matrix can be written in compact form as:

$$V_{S_{Full}} = \begin{bmatrix} V_{S_{SNP}} & \\ 0 & V_{S_{LDSC}} \end{bmatrix}$$

Stage 2 Estimation

In Stage 2, the genetic covariance matrix (*S*) obtained in the previous stage, is used to estimate the parameters in a SEM. The fit function minimized in the diagonally weighted version of WLS estimation that is standard in the Genomic SEM software package is the following:

$$F_{WLS}(\theta) = (s - \sigma(\theta))' D_S^{-1} (s - \sigma(\theta)),$$

where *S* and Σ(θ) have been half-vectorized to produce s and σ(θ) respectively, and D_S is V_S with its off-diagonal elements set to 0. We choose the diagonally weighted version of WLS because it is more tractable to implement for large (highly multivariate) matrices and is more stable than fully weighted WLS in finite samples (Flora & Curran, 2004; Savalei, 2014).

The WLS fit function will produce consistent estimates of the model parameters when the model is true (Savalei, 2014). However, the "naïve" *SE*s and fit statistic produced in Stage 2 estimation will be incorrect, because the estimator does not use the full V_S matrix in estimation. Thus, robust corrections must be applied to produce consistent estimates of *SE*s and test statistics. The correct sampling covariance matrix of the Stage Two, Genomic SEM parameter estimates (i.e., V_θ) can be obtained using a sandwich correction (Savalei, 2014; Savalei & Bentler, 2009):

$$V_\theta = \left(\hat{\Delta}'\Gamma^{-1}\hat{\Delta}\right)^{-1} \hat{\Delta}'\Gamma^{-1} V_S \Gamma^{-1}\hat{\Delta}\left(\hat{\Delta}'\Gamma^{-1}\hat{\Delta}\right)^{-1}$$

where $\tilde{\Delta} = \left.\frac{\partial LD(\theta)}{\partial \theta'}\right|_{\theta=\tilde{\theta}}$ is the matrix of model derivatives evaluated at the parameter estimates, and G is the naïve Stage 2 weight matrix.

Q_{SNP} Test of Heterogeneity

As part of Genomic SEM, an index is provided that quantifies the extent to which an observed vector—consisting of univariate regression effects of a given SNP on each of the phenotypes—can be explained by a *common pathway* model that assumes that the effects on the phenotypes are entirely mediated by the common genetic factor(s). In other words, this index enables the identification of loci that do and do not operate on the individual phenotypes exclusively by way of their associations with the common factor(s). Because of its intuitive and mathematical similarity to the meta-analytic Q-statistic used in standard meta-analyses to index heterogeneity of effect sizes (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006) we label this heterogeneity statistic, Q_{SNP}. Q_{SNP} is a χ^2-distributed test statistic with larger values indexing a violation of the null hypothesis that the SNP acts entirely through the common factor(s).

Q_{SNP} is calculated using a two-step procedure. In Step 1, a *common pathway* model is fit in which both factor loadings, the SNP effect on the common factor(s), and the residual variances of the common and unique factors are freely estimated (with one factor loading fixed to unity for factor identification and scaling). No paths representing direct effects of the SNP on the genetic components of the individual phenotypes are estimated. In Step 2, a *common plus independent pathways model* is specified, in which the factor loadings and the SNP effect on the

common factor are fixed to the values estimated in Step 1, and direct effects of the SNP on individual indicators and the residual variances of each indicator are freely estimated.

METHOD

Standardization and Scaling of Summary Statistics for Multivariate GWAS

Typically, GWAS summary statistics for quantitative phenotypes are not reported in terms of covariances but as ordinary least squared (OLS) unstandardized regression coefficients, with the phenotypes standardized prior to analyses (i.e., the coefficients are standardized with respect to the outcome, but not the predictor). In order to transform these partially standardized regression coefficient ($b_{SNP,P}$) of a SNP effect on phenotype P to a covariance, we multiply by the variance of scores on the SNP. The variance (s^2_{SNP}) of scores (0, 1, 2) of a biallelic autosomal SNP is estimated as 2pq, assuming Hardy-Weinberg-Equilibrium, where p = the minor allele frequency (MAF) and q = 1-MAF, with the MAF typically obtained from a reference sample. As the latent genetic factors estimated in LDSC are scaled relative to unit-variance scaled phenotypes (by virtue of the SNP heritability estimates being placed on the diagonal of *S*), no further scaling is needed to transform this SNP-phenotype covariance into a SNP-genotype covariance.

When OLS regression coefficients and standard errors are provided from an analysis in which the phenotype has not been standardized prior to analyses, or only Z statistics or *p*-values (for which Z statistics can be readily obtained) are provided, the partially standardized regression coefficients and their standard errors can be obtained as $Z = \frac{b^*_{SNP,P}}{SE_{b^*_{SNP,P}}}$, $b_{SNP,P} = \frac{Z}{\sqrt{N\sigma^2_{SNP}}}$, and

$SE_{b_{SNP,P}} = \frac{b_{SNP,P}}{Z}$, where $b^*_{SNP,P}$ is equal to the regression coefficient for the OLS GWAS of the unstandardized phenotype. These derived partially standardized coefficients are then transformed into covariances by multiplying by the variance of scores on the SNP, as per above.

When the GWAS summary statistics are reported for logistic regressions of liabilities for categorical outcomes (e.g. case/control status) on the SNP, the logistic regression coefficients can be transformed into covariances as above, by multiplying by the SNP variances. However, it is appropriate to further transform the coefficients and their *SE*s such that they are scaled relative to unit-variance scaled liability. This can be achieved by dividing by $\sqrt{s^2_{SNP} \times b\text{logit}^2_{SNP,P} + \frac{p^2}{3}}$, as a logistic regression model implies a residual variance of $\frac{p^2}{3}$. If GWAS summary statistics are reported for odds ratios (ORs), they can be transformed to logistic regression coefficients by taking their natural logarithm. Standard errors for the logistic regression coefficient are obtained as SE_{OR}/OR. The derived logistic coefficients and their *SE*s should further be transformed such that they are scaled relative to unit-variance scaled phenotypes, as per above. Note that when the outcomes are categorical, the liability scale heritabilities and genetic covariances from multivariable LDSC (and not what are referred to as the "observed scale" heritabilities and genetic covariances) should be used to populate the *S* matrix. This has the desirable property of both modeling the continuous scale of risk in the population and providing estimates that are independent of the observed prevalence of the categorical outcomes.

On occasion, summary statistics will be provided from OLS GWASs of categorical outcomes (e.g., case/control status). Such an analysis is sometimes referred to as a linear

probability model, as it (incorrectly) assumes that the association between the predictor and the probability of being in the comparison (e.g. case) group relative to the reference (e.g. control) group is linear. Parameters from the linear probability model are dependent not only on the strength of the association between the SNP and the continuous underlying liability, but also on the MAF and the proportion of comparison group members (cases) in the sample. Thus, parameters from the linear probability model cannot be used directly in Genomic SEM. However, particularly in the case of complex traits, for which the effect sizes for individual SNPs are small, results from the linear probability model can be used to very closely approximate logistic regression coefficients and *SE*s that are amenable for use in Genomic SEM (Lloyd-Jones, Robinson, Yang, & Visscher, 2018). This approximation can be obtained as $Z = \frac{b^{**}_{SNP,P}}{SE_{b^{**}_{SNP,P}}}$, $b\text{logit}^{*}_{SNP,P} = \frac{Z}{\sqrt{v(1-v)Ns^{2}_{SNP}}}$, and $SE_{b_{SNP,P}} = \frac{b\text{logit}^{*}_{SNP,P}}{Z}$, where $b^{**}_{SNP,P}$ is equal to the regression coefficient from the linear probability model, $b\text{logit}^{*}_{SNP,P}$ is the expected logistic regression coefficient that is derived from the linear probability model results, *v* is equal to the proportion of cases in the sample, and s^{2}_{SNP} is the variance of the SNP, computed from its MAF obtained from a reference sample, as per above. To scale the derived logistic coefficient such that it is scaled relative to unit-variance scaled liability, the coefficient should be divided by $\sqrt{s^{2}_{SNP} \times (b\text{logit}^{*}_{SNP,P})^{2} + \frac{p^{2}}{3}}$, as is done in the case of logistic regression (case/control) effects described above.

Model Fit Statistics

Model χ^2 is an index of exact fit of a SEM. It indexes whether the model-implied genetic covariance matrix, $\Sigma(\theta)$, differs from the empirical genetic covariance matrix, *S*. Model χ^2 can also be used as a relative fit index for comparing nested models. Conventional SEM approaches to indexing model χ^2 are based on formulas that directly incorporate *N*. Because there is not an *N* that directly corresponds to the genetic covariance matrix that is modelled by Genomic SEM in the same way that *N* typically corresponds to an observed covariance matrix, a formula was derived for estimating model χ^2 that does not require *N*, but instead incorporates the sampling covariance matrix of the model residuals. This is done in two steps. In Step 1, the proposed model (e.g., a common factor model) is estimated. In Step 2, all of the Step 1 estimates are fixed, and the residual covariances and residual variances of the indicators are freely estimated. Residual variances are estimated in Step 2 by estimating the variances of *k* residual factors defined by the indicators. This provides an estimate of the discrepancy between the model implied and observed covariance matrices, $R = S - \Sigma(q)$, along with the sampling covariance matrix (V_R) of *R*. While the discrepancy between model implied and observed covariance matrices can be computed simply by deriving covariance expectations from the Step 1 model and subtracting the observed covariance matrix, such an approach would not provide the corresponding V_R matrix necessary for the calculations below. The V_R matrix is expected to be positive semidefinite and, consequently, have no negative eigenvalues. Therefore, the V_R matrix has the following eigendecomposition:

$$V_R = (P_1\ P_0) \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1' \\ P_o' \end{pmatrix}$$

where P_1 is a matrix of principal components (eigenvectors) of V_R, and E is a corresponding diagonal matrix consisting of non-zero eigenvalues. P_0 reflects the null space of V_R. Projecting R_i—a vector of residual covariances estimated from the Step 2 Model—onto P_1 and adjusting for corresponding eigenvalues, we have that:

$$E^{\frac{-1}{2}} P_1' R_i N(0, I_r)$$

Therefore,

$$R_i' P_1 E^{-1} P_1' R_i \sim \chi^2 (df)$$

This equation produces a test statistic that is χ^2 distributed with degrees of freedom equal to the difference between the number of nonredundant elements (k*) in the empirical covariance matrix (S) and the number of freely estimated parameters in the proposed model, as would be the case in any regular SEM.

The Comparative Fit Index (CFI) is a test of approximate model fit. CFI indexes the extent to which the proposed model fits better than a model that allows all phenotypes to be heritable, but assumes that they are genetically uncorrelated. The χ^2 statistic can be used to calculate CFI by calculating a second χ^2 statistic for a so-called independence model, i.e. a model that estimates genetic variances of all phenotypes but assumes all genetic covariances to be zero, such that $\sum(\theta)$ is diagonal. CFI is calculated using the formula below (Kenny, 2014), with $f = \chi^2$ – degrees of freedom:

$$\frac{f(\text{Independence Model}) - f(\text{Proposed Model})}{f(\text{Independence Model})}$$

For the χ^2 of the independence model, a model is estimated in Step 1 that includes only the variance of the indicators and no common factor. In Step 2, these variances are fixed and the covariances among the indicators and variances of *k* residual factors defined by the indicators are estimated and used to populate the same equation above used to calculate the proposed model χ^2. CFI values theoretically range from 0 to 1, with higher values indicating good fit. CFI values of .90 and above are typically considered acceptable fit, and values of .95 and above are typically considered good model fit (Kaplan, 2008). When the empirical covariance matrix contains a large number of cells that are very close to 0, CFI values may be low, even when such cells are approximated well by the model.

Akaike Information Criterion (AIC) is a relative fit index that balances fit with parsimony, and can be used to compare models regardless of whether they are nested. AIC is calculated as:

$$AIC = \chi^2 + 2 \times fp,$$

where fp is the number of free parameters in the model (Tanaka, 1993). Lower AIC values are considered superior.

Standardized Room Mean Square Residual (SRMR) is an index of approximate model fit that is calculated as the standardized root mean squared difference between the model-implied and observed correlations in $\Sigma(\theta)$ and *S*, respectively (Hu & Bentler, 1999). Higher SRMR values indicate a larger discrepancy between $\Sigma(\theta)$ and *S*. It is positively-biased, with larger bias resulting when the contributing univariate GWAS samples are lower powered. SRMR values below .10 indicate acceptable fit, values less than .05 indicate good fit, and a value of 0 indicates perfect fit (Bentler & Hu, 1995).

Simulation of Factor Structure

In order to evaluate the ability of Genomic SEM to capture the genetic factor structure in the generating population, the GCTA package (Yang et al., 2011) was used to generate 100 sets of 6 independent, 100% heritable phenotypes ("orthogonal genotypes") to pair with genotypic data for 39,909 randomly selected, unrelated individuals of European descent from UKB data for the 1,209,498 SNPs present in HapMap3. The generating list of causal SNPs was set to 10,000 for all 600 genotypes, with the specific list of causal variants sampled with replacement from the 1,209,498 SNPs. One of the six orthogonal genotypes per set was designated an index of the general genetic factor and the remaining five were designated indices of domain-specific genetic factors. All of these orthogonal genotypes were scaled to *M*=0, *SD*=1. Five new correlated genotypes were then constructed, each as the weighted linear combination of the general genetic factor and one domain-specific genetic factor. Weights for contribution of the general genetic factor were $\lambda_{Fg,k}$ =.70, .60, .50, .40, and .30, for correlated genotypes 1-5, respectively. Weights for the domain-specific factors were $\sqrt{(1 - \lambda_{Fg,k}^2)}$. Phenotypes were then each constructed as the weighted linear combination of one of the correlated genotypes and domain-specific environmental factors (randomly sampled from a normal distribution with *M*=0, *SD*=1). Heritabilities for phenotypes 1-5 were set to h_k^2=35%, 40%, 50%, 60%, and 70%, respectively, such that the weights for the genotypes were $\sqrt{h_k^2}$ and the weights for the environmental factors were $\sqrt{(1 - h_k^2)}$. We chose these figures to stabilize the properties of the distributions across simulations at 100 replications with *N*~39K each. We expect that with lower SNP h^2's, the same patterns would hold, albeit at larger sample sizes. Each of the 500 phenotypes (100 sets of 5

phenotypes) was then analyzed as a univariate GWAS in PLINK (Purcell et al., 2007) to produce univariate GWAS summary statistics. Our multivariable LDSC function was then used to construct 100 sets of 5×5 genetic covariance matrices (*S*) and associated sampling covariance matrices (V_S), and Genomic SEM was used to fit a one factor model to each set.

Simulation of Partial Sample Overlap

In order to examine the effect of sample overlap on estimates obtained from Genomic SEM, the GCTA package (Yang et al., 2011) was used to generate a 50% heritable, quantitative phenotype with 30,000 causal SNPs. The phenotype was paired with genetic data from 100,000 randomly selected, unrelated individuals of European descent from UKB data for 1,209,498 HapMap3 SNPs. Three sets of 60,000 participants each were created using this same phenotype, with 40,000 participants overlapping across all three identical phenotypes and 20,000 participants unique to each phenotype (i.e., 100,000 total participants). These three subsamples were individually analyzed in PLINK (Purcell et al., 2007) to produce univariate GWAS summary statistics. The multivariable LDSC function was then used to construct the genetic covariance and sampling covariance matrix using the three sets of summary statistics, and Genomic SEM was used to fit a one factor model with the SNP predicting the common factor. Two key pieces were verified at this stage. First, we confirmed that the standardized factor loadings on the common factor were 1 for the identical phenotypes. Second, we verified that the bivariate LD-score intercepts that are used to account for sample overlap in the sampling covariance matrix were as expected. The equation for the LD-score bivariate intercept is: $N_s\rho/\sqrt{(N_1N_2)}$, where N_s = sample overlap, ρ = the phenotypic correlation, N_1 = sample size of trait 1, and N_2 = sample size of trait 2. In this simulation, we observed bivariate intercepts of .67,

which is as expected given sample overlap of 40,000, a phenotypic correlation of 1, and sample sizes of 60,000 (i.e., 40,000*1/√(60,000*60,000) = .67). Finally, estimates from this multivariate GWAS were compared to estimates from the univariate GWAS in PLINK for the full set of 100,000 participants.

RESULTS

Validation of Summary-Based Model Fit Statistics via Simulation

A generating population with a common factor model defined by four, five, or six indicators was used to examine the null distribution of the newly derived χ^2 test statistic using a set of 1,000 simulations per model. These simulations did not include individual genotypes, and were simulated solely based on a generating factor structure. For the six indicator models the standardized factor loadings in the generating population were .42, .64, .22, .59, .19, and .64. The four and five indicator models specified the same factor loadings, excluding the last, or last two loadings, respectively. Results indicated that the two-step procedure described above produced a test statistic equivalent to the χ^2 statistic calculated by lavaan from the raw data ($r > .99$). For a χ^2 distributed test-statistic, the mean of the null sampling distribution should match the *df* of the test. As expected, the distribution of the test-statistic conformed to a χ^2 distribution with an average approaching the *df*. Calculated CFI values were also highly consistent ($r > .99$) with those observed using the CFI statistic provided by lavaan when using raw data. Calculated AIC values were not contrasted with those obtained using the lavaan package in R in the simulations below as the software uses a formula that includes a log-likelihood estimate contingent on the provided sample size.

Validation of Null Distribution of Q_{SNP}

To verify that the null distribution for Q_{SNP} is χ^2 distributed, a set of simulations specified a generating population in which the direct effects of the SNP on the indicators were entirely mediated through the common factor. Each simulation included 1,000 datasets, with N = 100,000 completely overlapping participants per dataset. We examined three models with F = 1 factor, and k = 4, 5, or 6 phenotypes. Using a genome-wide significance threshold, in all cases the false discovery rate for Q_{SNP} was 0, and the power to detect a SNP effect on the common factor was 1. Mean estimates of Q_{SNP} were approximately equal to the *df* of the corresponding model. Figure 1 depicts the null sampling distributions of Q_{SNP}. These results indicate that Q_{SNP} is approximately χ^2 distributed.

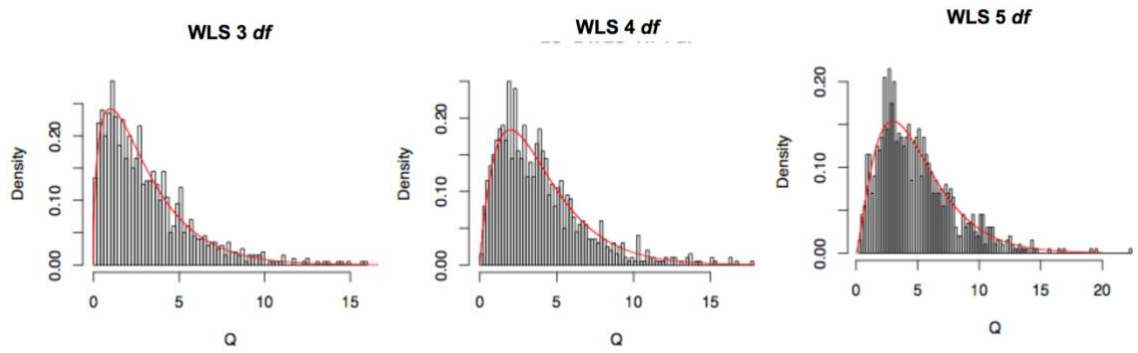

Figure 1. Null distributions of Q_{SNP} for 1,000 simulations Per Model. Red lines for all panels depict the chi-square distribution with the relevant *df*. The top, middle, and bottom panels depict the sampling distributions for 3, 4, and 5 *df*, respectively.

Recovery of Model Population Parameters

We performed 100 runs of Genomic SEM on raw individual-level genotype data for which we simulated multivariate phenotypic data to conform to a single genetic factor model (a latent trait that partially causes 5 observed outcomes). Across the 100 simulations Genomic SEM estimates closely matched the parameters specified in the generating population (Figure 2).

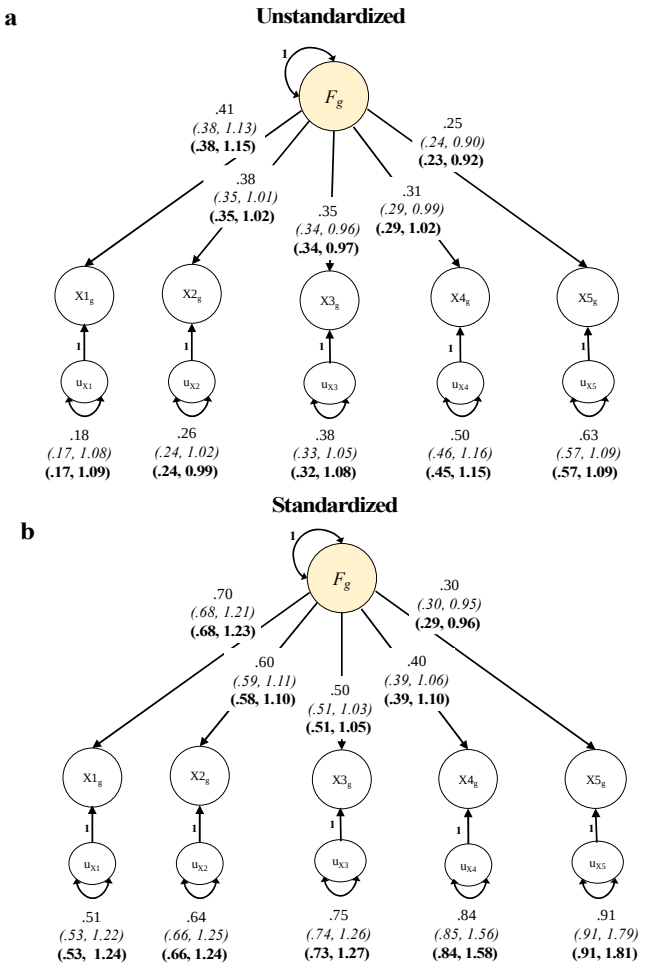

Figure 2. Genomic SEM Simulation Results. Results from 100 runs of Genomic SEM using data simulated at the level of the SNPs. Results are presented for unstandardized (panel a) and standardized (panel b) estimates. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). The ratio of mean and empirical *SE*s was close to 1 in all cases, although slightly above 1 (i.e., conservative) for standardized estimates of residual variance. These *SE estimates* are expected to be upwardly biased in the standardized case due to genetic variance estimates being rescaled to exactly 100%.

Model *SE*s also closely matched the standard deviations of parameter estimates. We also compared fit statistics (CFI, AIC, and model χ^2) for the correctly specified common factor model and two deliberately misspecified models: (i) a model in which all indicators were constrained to have the same factor loading, and (ii) a model for which the loading of the third indicator was set to 0. As expected, results indicated that the common factor model matching the generating population was favored $\geq$ 99% of the time across model fit indices (Figure 3).

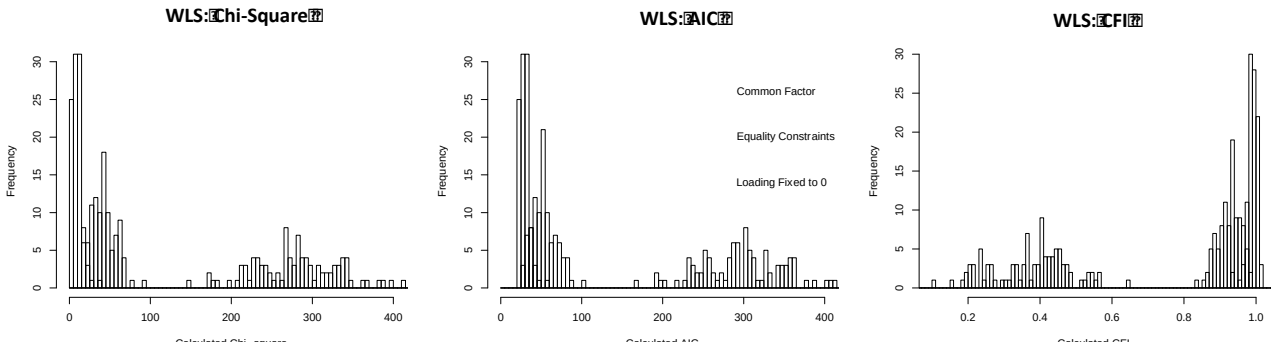

Figure 3. Model Fit Indices from Genomic SEM Simulations. Model fit indices were compared across the 100 runs of Genomic SEM using simulated data. Depicted in blue are model fit indices for runs specified to match the generating population (i.e., one common factor with freely estimated factor loadings). Depicted in green are indices for models specified to have equal factor loadings across all indicators. Depicted in red are indices for a model in which the third indicator loading was fixed to 0. Indices favored the model that matched the generating population for model chi-square, AIC, and CFI in 100% of cases, with the exception that 99 models favored the matching model for AIC.

Simulation of Sample Overlap

One major benefit of Genomic SEM is that summary statistics can be used from samples with unknown degrees of sample overlap. We performed a simulation in order to verify that the inclusion of data from overlapping samples does not bias Genomic SEM parameter estimates or their standard errors. We simulated data for a single quantitative phenotype in 100,000 participants, and subsequently split the sample into three subsamples of 60,000 participants each

(with ~66% pairwise overlap between the subsamples). We submitted each of the three subsamples to an independent GWAS, and used the three resulting sets of summary statistics as input for a Genomic SEM model, in which we specified the phenotype from each individual set of summary statistics as a different indicator of a common factor, onto which a SNP effect was specified. If sample overlap is *not* appropriately accounted for, data are incorrectly treated as deriving from 180,000 participants (as opposed to 100,000 total participants), and we would expect the Z-statistics for the SNP effects to be upwardly biased relative to those from a univariate GWAS applied directly to the single phenotype in the 100,000 participants. As expected, we observed no such bias. A linear regression of estimates from the Genomic SEM model from the three overlapping samples of 60,000 participants each predicting univariate GWAS Z statistics in the complete sample of 100,000 participants revealed near perfect correspondence (unstandardized slope = 1.003, intercept = -.003).

DISCUSSION

Applications of genome-wide methods to data from large scale population-based samples have uncovered clear evidence of pervasive shared genetic architecture across social, behavioral, and psychiatric traits. We created the Genomic SEM framework for modeling the multivariate genetic architecture of constellations of genetically correlated traits and incorporating genetic covariance structure into multivariate GWAS discovery. In contrast to existing methods (Verhulst et al., 2017) that model phenotypic, rather than genetic covariance structure, and rely on raw data, Genomic SEM employs summary GWAS data to model genetic covariance structure. Genomic SEM is computationally efficient, accounts for potentially unknown degrees

of sample overlap, and allows for flexible specification of covariance structure, such that several broad classes of structured covariance models can be applied.

The Genomic SEM approach shares benefits of some existing approaches (Turley et al., 2018) for boosting power by combining information across genetically correlated phenotypes. However, Genomic SEM uniquely allows one to compare different hypothesized genetic covariance architectures and to incorporate such architectures into multivariate discovery. Importantly, shared genetic liabilities across phenotypes can be explicitly modeled as factors that may be treated as broad genetic risk factors with equally broad downstream consequences. Multivariate genetic methods have existed for decades in the twin literature, with Martin and Eaves (1977) providing a framework for fitting structural equation models of genetic and environmental variance components in multivariate twin data. Genomic SEM can be used to reproduce multivariate genetic models from the existing twin literature using GWAS summary data from unrelated individuals. Moreover, Genomic SEM offers new promise as a method that allows for the estimation and modeling of genetic covariance even among phenotypes for which phenotypic covariance cannot be estimated.

Genomic SEM is not the only method for multivariate GWAS. Other methods, such as MTAG (Turley et al., 2018), SHom/SHet (Zhu et al., 2015), metaUSTAT (Ray & Boehnke, 2018), min-P (O'Reilly et al., 2012), and TATES (Van der Sluis, Posthuma, & Dolan, 2013) allow researchers to perform multivariate meta-analyses based solely on summary data. The methods can generally be divided into 2 distinct classes: methods which aggregate test statistics or effect sizes based on a model (Genomic SEM, SHom and MTAG) and those which efficiently select from the univariate *p*-values while taking care not to inflate Type-I error (min-P, TATES,

and SHet). As we show with respect to MTAG, some models can be fit in the context of Genomic SEM. It is also conceivable that the approach of efficiently selecting the minimum *p*-value from a set of analyses while maintaining proper Type-I error control could be integrated within Genomic SEM. For instance, whereas TATES is currently applied to select the minimum *p*-value from a series of univariate analyses of correlated traits, the same analysis could be used to select the minimum *p*-value from a series of Genomic SEM models.

In contrast to approaches that assume homogeneity of effects across SNPs, such as MTAG (Turley et al., 2018), Genomic SEM includes diagnostic indices for its key assumptions, including a test for heterogeneity, Q_{SNP}, that can be applied at the level of the individual SNPs. This offers the unique ability to identify SNPs that confer specific risk to individual phenotypes, symptoms, or indicators. This question may be of particular interest as the large degrees of genetic overlap identified across phenotypes (e.g., bipolar disorder and schizophrenia) beg the question: what are the genetic causes of phenotypic divergence? Whereas previous GWASs have combined items tapping genetically-related phenotypes into a single score, or even combined cases with different diagnoses to obtain a shared genetic effect, Genomic SEM allows researchers to interrogate shared genetic effects between diagnoses or indicators, while concurrently testing for causes of divergence (i.e., loci that are related only to a specific phenotype, or subset of phenotypes, but not the more general liability). Because Genomic SEM relies only on GWAS summary data, it can be applied to a broad spectrum of traits, including social, economic, cognitive, and, of most relevance to the current dissertation, psychiatric outcomes.

Chapter 2: Validating Stratified Genomic SEM[2]

INTRODUCTION

In light of high levels of polygenicity for psychiatric traits (Hyman, 2018), it is unclear what stratum of genomic measurement is most useful. At the level of individual genetic variants, there are unlikely to be a handful of key loci that explain significant variation in comorbidity patterns. Zooming out to genome-wide estimates of genetic correlations is valuable with respect to providing an overarching sense of overlap, but this wide lens provides limited biological insight. Univariate methods for biological annotation, such as Stratified LD-score Regression (S-LDSC; Finucane et al., 2015) offer a useful middle ground between individual loci and genome-wide estimates by identifying genetic enrichment within functional annotations (e.g., genes expressed in the hippocampus; evolutionarily conserved genes). However, these approaches are currently restricted to the univariate case, and it would be inappropriate to infer enrichment of genetic covariance based solely on convergent univariate findings. For example, two traits may be highly enriched within a specific annotation, but the genetic covariance in this annotation may still be near 0 if the two traits are affected by a small number of disparate loci with large effects.

Stratified Genomic SEM is an extension of the existing Genomic SEM framework outlined in Chapter 1. Broadly speaking, Stratified Genomic SEM fits structural models to genetic covariance matrices estimated within (potentially overlapping) biological categories.

---

[2] Results reported in the second chapter appear in the published work (Grotzinger et al., 2019). My contributions to this work included QC and gathering of the datasets, analyses, and write-up of the results. Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., ... & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature human behaviour*, *3*(5), 513-525.

These categories are constructed based on collateral gene expression data, such as that obtained from single-cell RNA sequencing, and from other bioinformatic annotations, such as allele frequency, and evolutionary conservation. Our multivariate extension of S-LDSC can be used to produce two types of genetic covariance matrices, S_0 and $S\tau$. S_0 is an estimate of genetic covariance within a specific biological category irrespective of overlap with other categories. Conversely, $S\tau$ provides an estimate of genetic covariance controlling for overlap. The difference between S_0 and $S\tau$ is similar to that between univariate and multiple regression coefficients, respectively.

The S_0 and $S\tau$ matrices can both be used as input to the same Two-Stage Structural Equation Modeling framework from traditional Genomic SEM. In Stratified Genomic SEM, Stage 1 consists of the estimation of the *stratified* genetic covariance matrix and corresponding *stratified* sampling covariance matrix. This sampling covariance matrix consists of squared *SE*s on the diagonal and sampling covariances on the off-diagonal, which capture dependencies between sampling errors that will arise due to situations such as sample overlap. These off-diagonals are estimated directly from our multivariate extension of S-LDSC, and thereby allow Stratified Genomic SEM to produce unbiased estimates for GWAS summary statistics with varying and unknown degrees of sample overlap. In Stage 2, an SEM is specified and parameters are estimated that reduce the discrepancy between the observed genetic covariance matrix (either S_0 or $S\tau$) and the model implied covariance matrix.

METHOD

Simulations of Stratified Genetic Covariance

The point estimates and standard errors (*SE*s) produced by Stratified Genomic SEM were evaluated using simulation. As partitions may overlap substantially, an additional goal was to determine whether Stratified Genomic SEM can produce unbiased estimates that are specific to the causal partition of interest. The expectation is that the zero-order genetic covariance matrices (S_0) will not show partition specificity due to the noted partition overlap. Conversely, the expectation is that the τ covariance matrices (Sτ)—which reflect the covariances estimated within a multiple regression framework—would demonstrate the desired specificity in estimates. The first step in the simulation procedure was to generate 100 sets of 45, 100% heritable phenotypes ("orthogonal genotypes") using the GCTA package (Yang et al., 2011). Each 100% heritable phenotype was specified to have 10,000 randomly selected causal variants from within a particular partition. These phenotypes were paired with genotypic data for 100,000 randomly selected, unrelated individuals of European descent from UKB data for the 1,209,498 SNPs present in HapMap3.

The simulated genotypes were used to construct six different factor structures for six causal partitions. All orthogonal genotypes were scaled $M=0$, $SD=1$. For three of the causal partitions (DHS Peaks, H3K27ac, and PromoterUSC) seven genotypes for each partition were used to construct six new correlated genotypes, each as the weighted linear combination of a domain-specific genetic factor and a general genetic factor, which was constructed from the seventh genotype. For the remaining three causal partitions (FetalDHS, H3K9ac, and TFBS) eight genotypes for each partition were used to construct two sets of three correlated genotypes for two correlated general genetic factors, constructed from the seventh and eighth genotypes. A set of six "total" genotypes was created by summing a factor indicator genotype from each of the

six causal partitions. As each genotype within each partition was specified to have 10,000 causal SNPs, the "total' genotypes created as the sum of six partition had 60,000 causal SNPs in the population generating model.

Phenotypes were subsequently constructed as the weighted linear combination of one of the six "total" genotypes and domain-specific environmental factors (randomly sampled from a normal distribution with *M*=0, *SD*=1). Heritabilities for phenotypes 1-6 were all set to h_k^2=60%, such that the weights for the genotypes were $\sqrt{h_k^2}$ and the weights for the environmental factors were $\sqrt{(1-h_k^2)}$. Each of the 600 phenotypes (100 sets of 6 phenotypes) was then analyzed as a univariate GWAS in PLINK (Purcell et al., 2007) to produce univariate GWAS summary statistics. The summary statistics were munged using the base *GenomicSEM* R package, and Stratified Genomic SEM coupled the 1000 Genomes Phase 3 BaselineLD Version 2.2 annotations was used to construct 100 sets of 6×6 stratified, zero-order genetic covariance matrices (S_0), τ covariance matrices (S_τ), and their corresponding sampling covariance matrices (V_{S_0} and V_{S_τ}).

RESULTS

Validating Stratified Zero-Order Genetic Covariances: S_0 and V_{S_0}

For the genetic covariance (S_0) matrix, the expectation is that the partition including all SNPs—the baseline (i.e., genome-wide) partition—will reflect the additive combination of the population generating covariance matrices specified for the six causal partitions. For the six causal partitions, the expectation is that the estimated covariance matrix will reflect a combination of the generating model in that partition plus the generating model in the other causal partitions weighted by the proportion of SNPs overlapping across the partitions. To test

these expectations, average observed covariance matrices were created across the 100 simulations for the genome-wide annotation and six causal annotations. The estimated genome-wide model approximately reflected an additive mixture of the six population generating covariance matrices, and was estimated with minimal bias (absolute value of mean discrepancy = .004; Figure 4). In addition, the observed covariance matrices within the causal partitions were, generally speaking, minimally biased relative to expectation (Figure 4; Table 1).

In order to evaluate *SE* performance, the ratio of the mean *SE* estimate was analyzed across the 100 simulations over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 simulations). A value above 1 for this ratio, therefore, indicates conservative *SE* estimates. This ratio was calculated within each of the partitions and for each cell of the covariance matrix. The average ratio across partitions and cells of the covariance matrix was 1.030 (Figure 5 for distribution across all partitions; Table 1 for ratio within causal partitions). Thus, the multivariate extension of S-LDSC produces a *SE* estimate for stratified heritability and covariance that performs as expected. In fact, these estimates are very slightly *conservative* as the mean *SE* was slightly larger than the empirical *SE*. Moreover, the average Z-statistic for heritability and covariance estimates within the causal partitions were all highly significant, suggesting more than adequate power when using the parameters of the current simulation (Table 1).

Table 1. Zero-order (S_0) and S_τ Covariance Results

Zero-Order Results					
Partition	**% HM3 SNPs**	***SE* Ratio**	**Absolute Mean Difference**	**Average h^2 Z-Statistic**	**Average Covariance |Z-statistic|**
Baseline	100%	0.956	0.004	36.99	24.85
DHS Peaks	16.80%	0.992	0.031	10.76	7.68
FetalDHS	12.95%	1.041	0.069	11.12	7.73
H3K27ac	45.10%	1.034	0.141	31.79	21.17
H3K9ac	16.20%	1.102	0.02	14.52	9.42
Promoter	5.65%	1.017	0.09	11.12	7.06
TFBS	16.86%	1.028	0.009	11.21	7.33
	τ Results				
Partition	**% Runs Producing Negative h^2**	***SE* Ratio**	**Absolute Mean Difference**	**Average h^2 Z-statistic**	**Average Covariance |Z-statistic|**
Baseline	100%	1.016	0.564	-2.64	-1.86
DHS Peaks	17%	0.988	0.008	1.74	1.51
FetalDHS	0%	1.034	0.006	2.85	1.98
H3K27ac	1%	1.034	0.013	2.62	1.85
H3K9ac	1%	1.043	0.006	2.81	1.57
Promoter	0%	1.034	0.007	5.72	3.52
TFBS	0%	1.034	0.003	2.99	1.72

Note. Average results reflect the mean across the 100 simulations described in the Chapter 2 Method section.

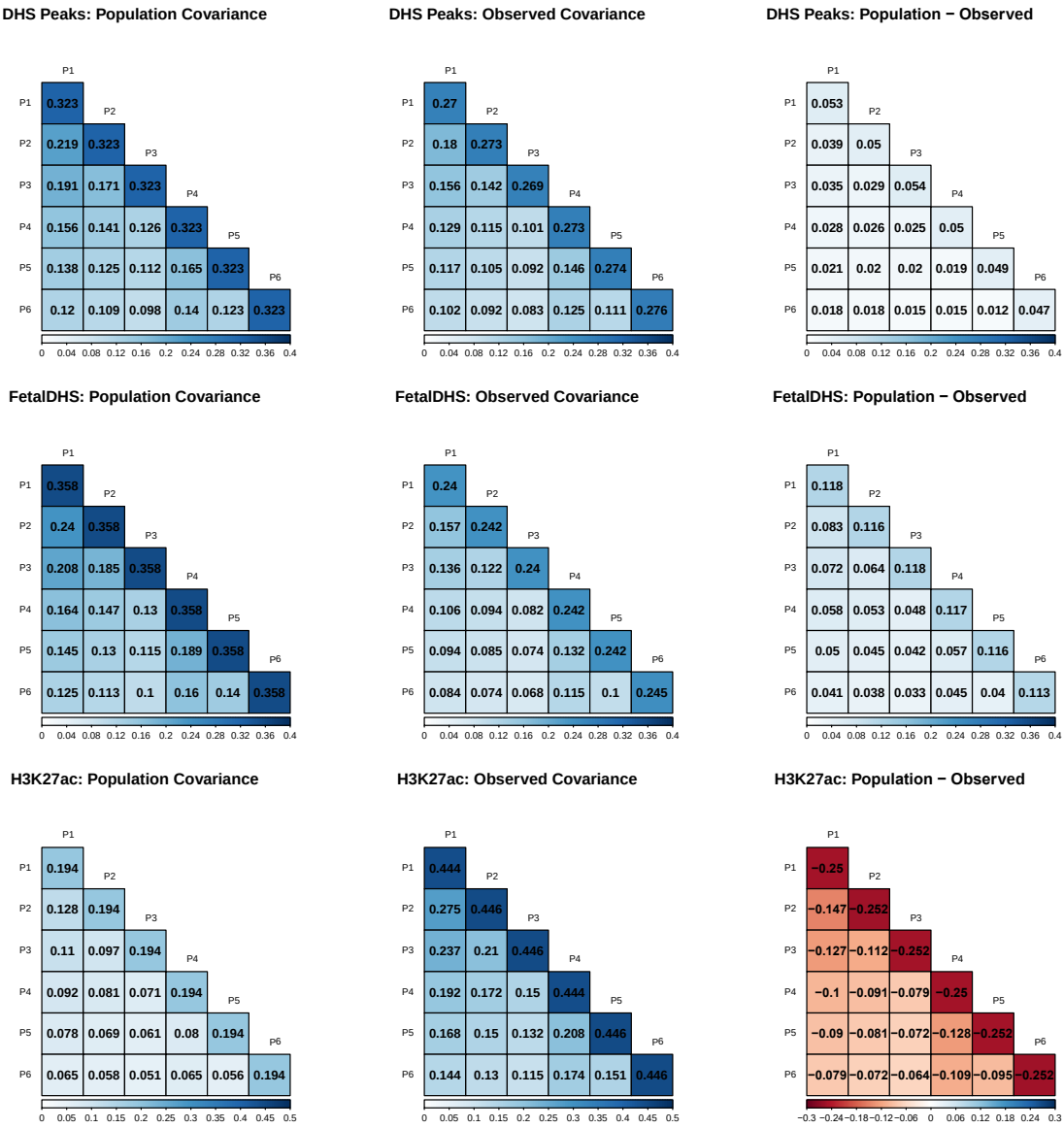

Figure 4a. Population Generating and Observed Zero-order (S_0) Covariance Matrices. The first column depicts the genetic covariance matrix in the generating population. The second column depicts the average observed covariance across the 100 simulations runs. The last column reflects the difference between the population matrix and average observed covariance matrix. For the zero-order matrices, estimates are expected to be biased in the sense that an individual partition will be affected by population generating covariances in overlapping partitions. Results are shown for the DHS Peaks (top row), Fetal DHS (middle row) and H3K27ac (bottom row) partitions.

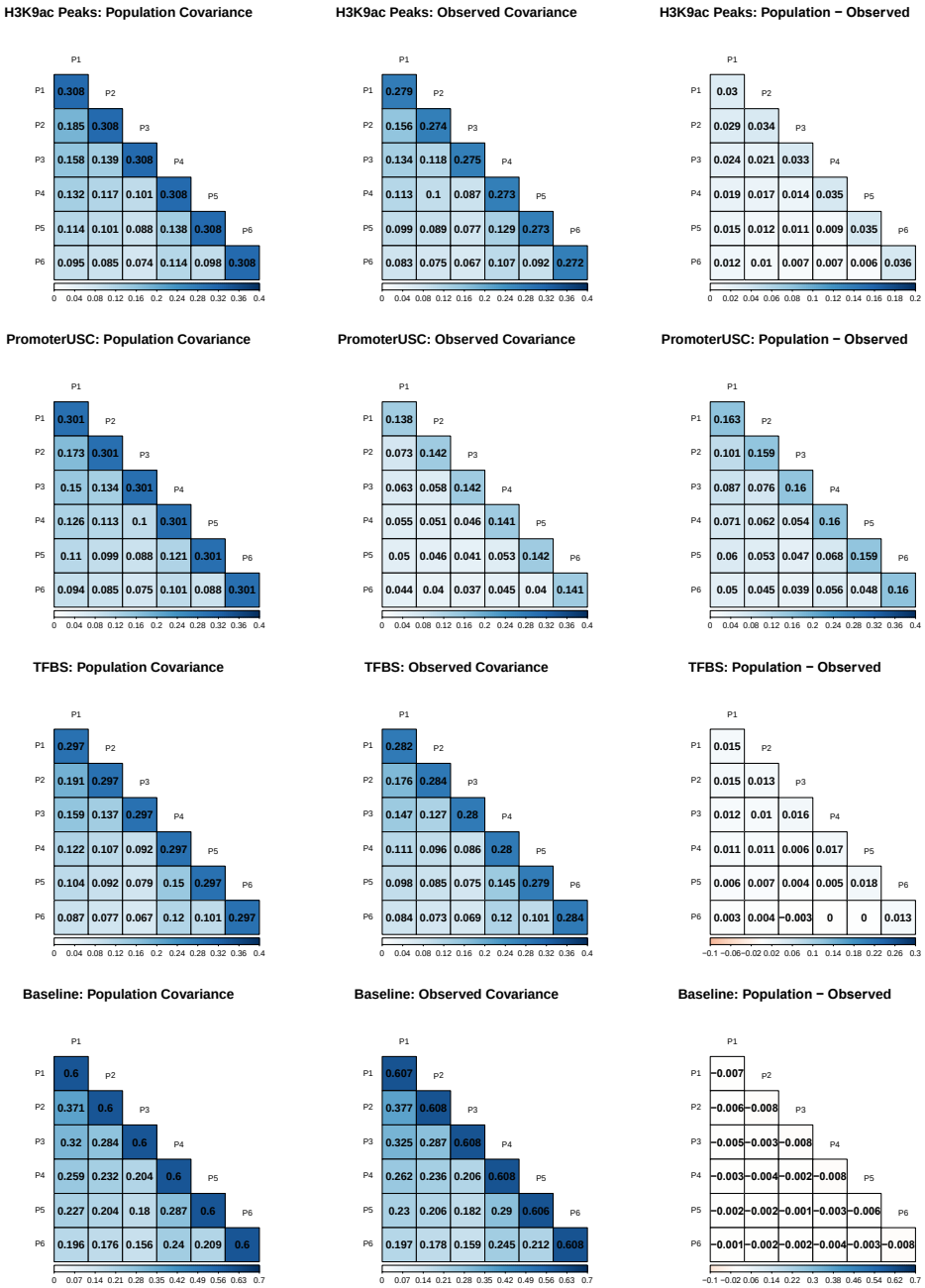

Figure 4b. Population Generating and Observed Zero-order (S_0) Covariance Matrices. The first column depicts the genetic covariance matrix in the generating population. The second column depicts the average observed covariance across the 100 simulations runs. The last column reflects the difference between the population matrix and average observed covariance matrix. For the zero-order matrices, estimates are expected to be biased in the sense that an individual partition will be affected by population generating covariances in overlapping partitions. Results are shown for the H3K9ac Peaks (top row), PromoterUSC (second row), TFBS (third row), and genome-wide (bottom row) partitions. Results for genome-wide estimates are unaffected by overlap with other partitions.

`

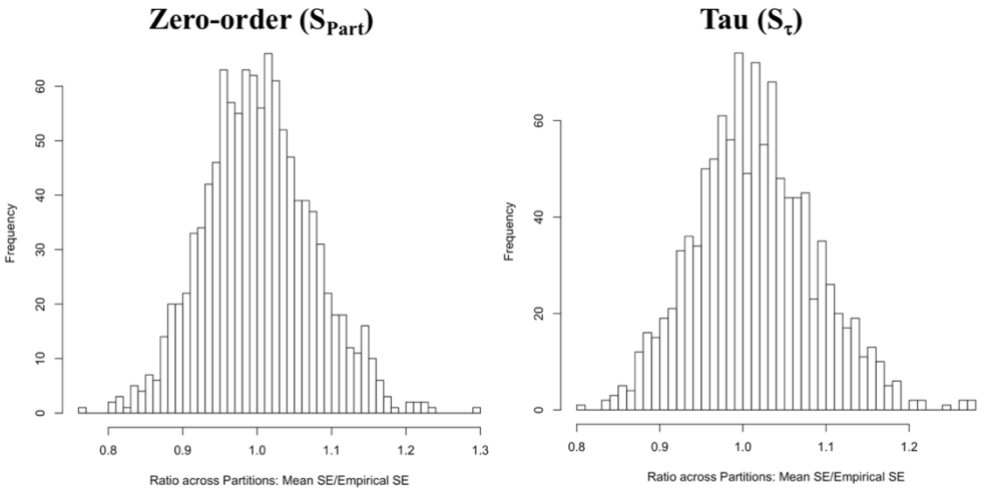

Figure 5. Distributions of *SE* Ratios. Panels depict mean *SE* ratios across the 100 simulations for the for mean *SE* over the empirical *SE* across the partitions for the S_τ (right panel) and zero-order (left panel) covariance matrices. Average ratios are shown for each cell of the covariance matrix for all partitions.

Validating Stratified τ Covariance : $S\tau$ and $V_{S\tau}$

The expectation for the genetic S_τ covariance matrices is that the observed covariance matrices will reflect the generating model within only that partition. Indeed, the causal partitions closely matched their respective population generating covariance matrices and bias was minimal (Table 1; Figure 6). The ratio of the mean *SE* estimate across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs) was subsequently evaluated. The average ratio of *SE* estimates was 1.014 across all partitions (Figure 5) and, importantly, was also close to 1 for the causal partitions (Table 1). Results for 4,459 of the total 5,300 covariance matrices produced negative heritability estimates. This included some of the causal partitions (Table 1), but was largely true for the non-causal partitions. Negative heritability estimates are unsurprising for the non-causal partitions as their population generating effect is 0 and will, therefore, form a sampling distribution centered

around 0 that includes negative values due to sampling variation. The Z-statistics for the τ heritabilities and covariances were, on average, smaller relative to the zero-order covariance matrices (Table 1). This is to be expected as the zero-order (S_0) covariance matrices include power gained from variance shared with overlapping partitions.

The $S\tau$ covariance matrices for the causal partitions were then used as input for Genomic SEM models. The two types of population generating models—a common factor and correlated factors model—were run for each partition. For all causal partitions, Genomic SEM estimates closely matched the parameters specified in the generating population (Figure 7). In addition, the ratio of the mean model *SE*s over the empirical *SE*s was near 1. Model fit statistics (CFI, AIC, and model χ^2) also generally favored the generating model for a particular partition (Table 2). This was least true for the H3K27ac partition. This is unsurprising as the population generating model for the H3K27ac partition—a correlated factors model with a factor correlation of .7—most closely matched the competing common factor model. Collectively, these results indicate that stratified Genomic SEM produces unbiased parameter estimates and standard errors for S_0 and $S\tau$, that $S\tau$ shows specificity to the causal partitions of interest, and that model fit indices generally favor the appropriate model.

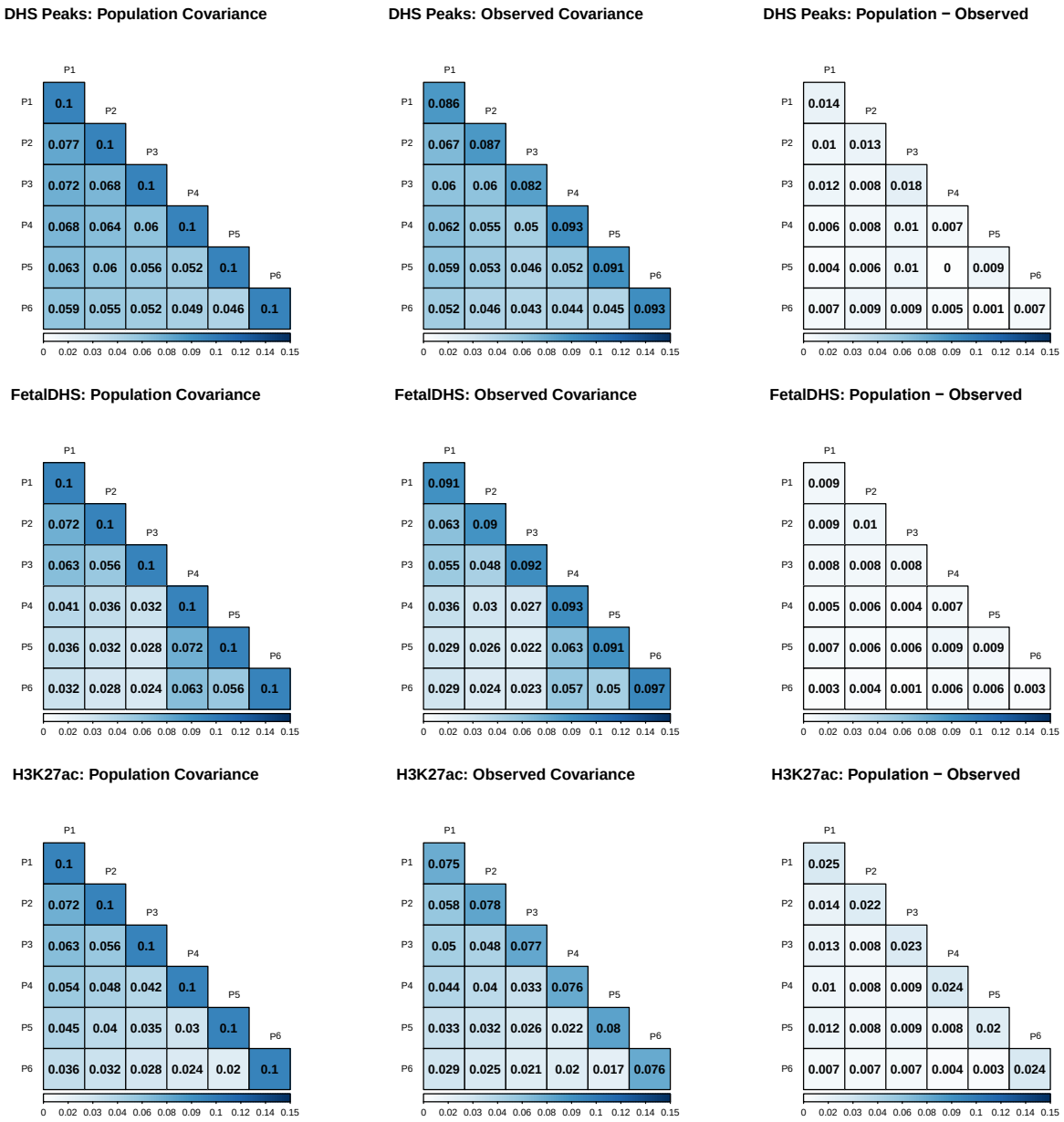

Figure 6a. Stratified S_τ Covariance Matrices. The first column depicts the genetic covariance matrix in the generating population. The second column depicts the average observed covariance across the 100 simulations runs. The last column reflects the difference between the population matrix and average observed covariance matrix. As observed, for the S_τ matrices, estimates are expected to be generally unbiased. Results are shown for the DHS Peaks (top row), Fetal DHS (middle row) and H3K27ac (bottom row) partitions.

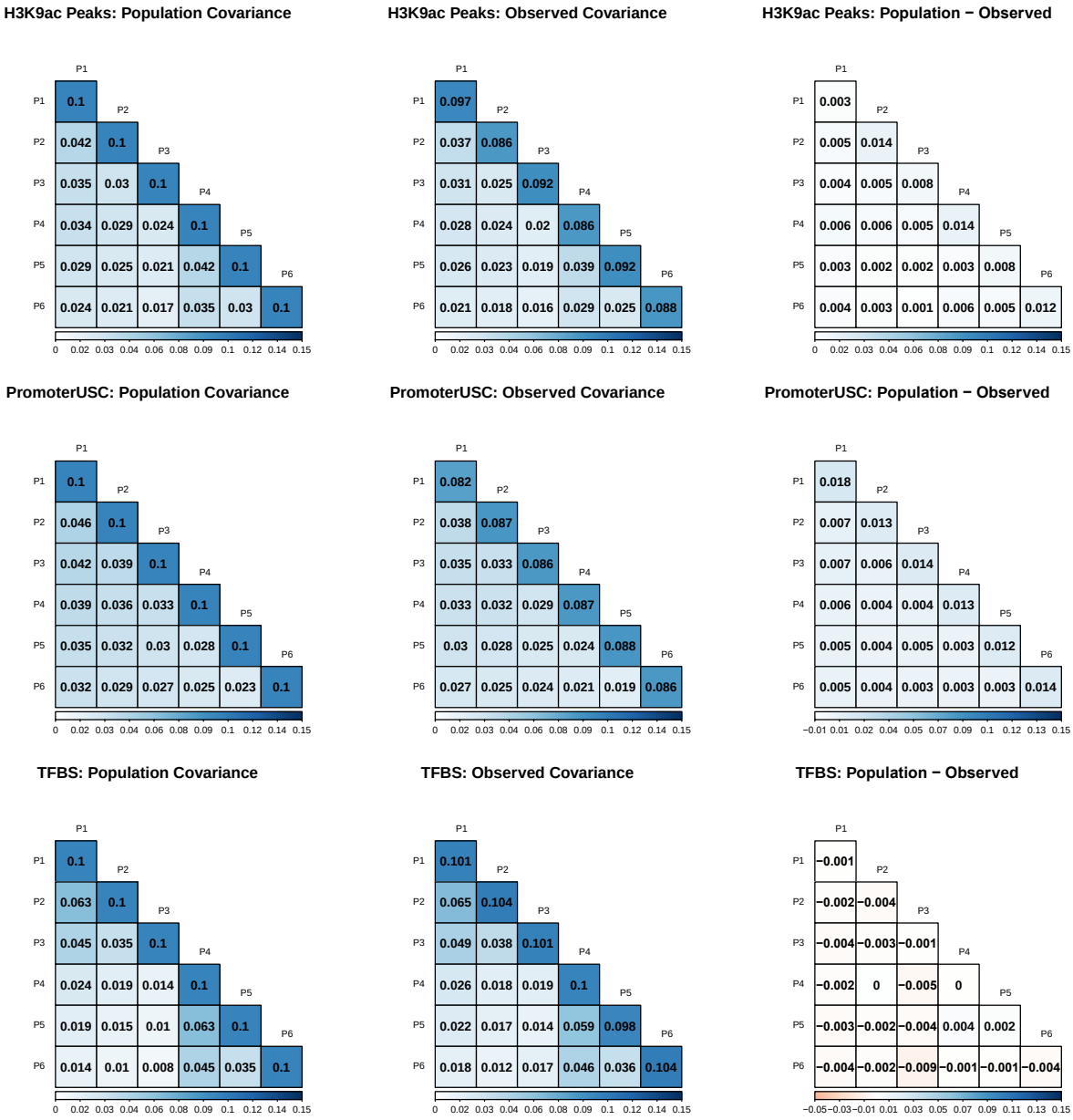

Figure 6b. Stratified S_τ Covariance Matrices. The first column depicts the genetic covariance matrix in the generating population. The second column depicts the average observed covariance across the 100 simulations runs. The last column reflects the difference between the population matrix and average observed covariance matrix. As observed, for the S_τ matrices, estimates are expected to be generally unbiased. Results are shown for the H3K9ac Peaks (top row), PromoterUSC (second row), TFBS (third row), and genome-wide (bottom row) partitions. Results for genome-wide estimates are unaffected by overlap with other partitions as the genome-wide partition includes all SNPs.

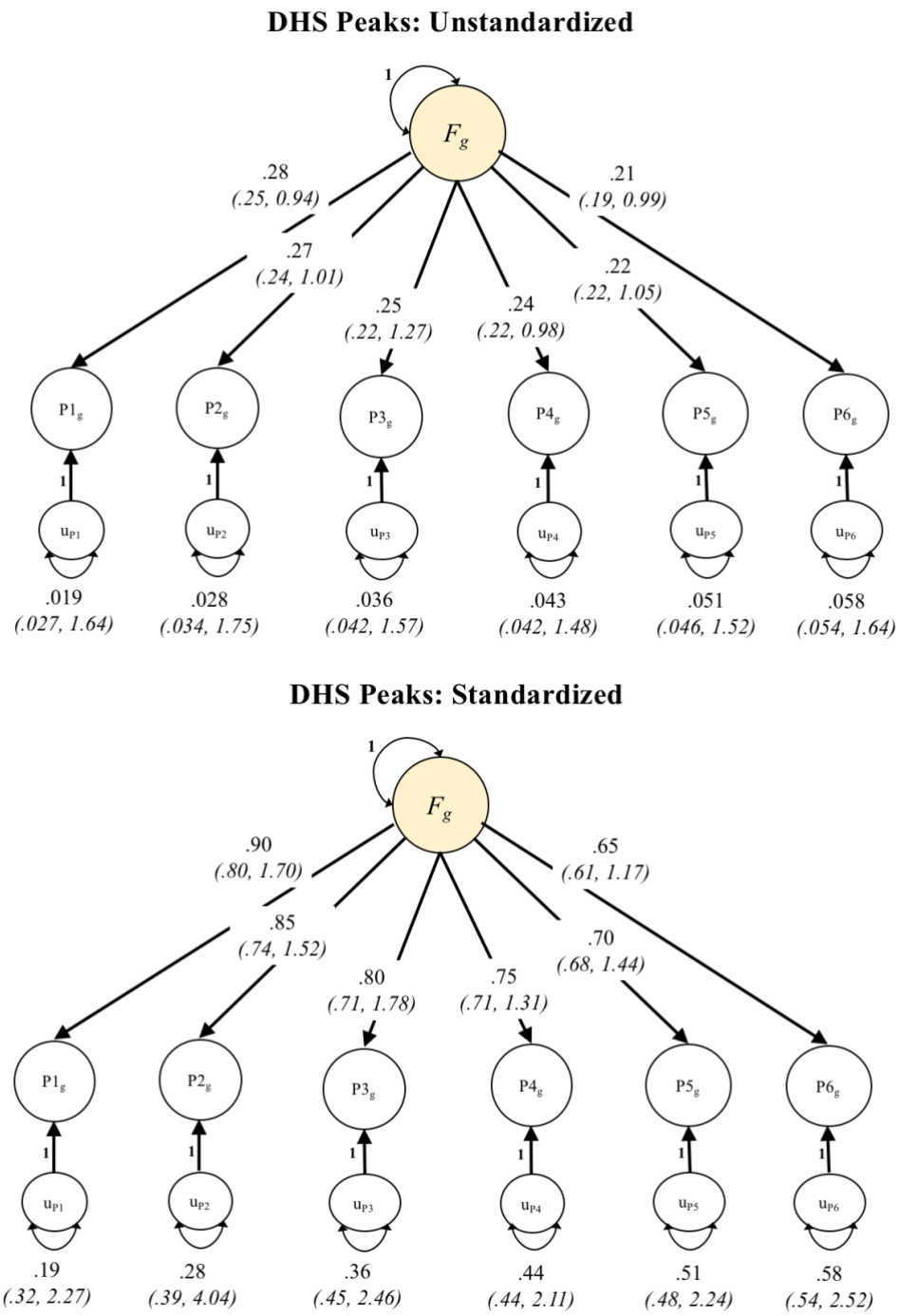

Figure 7a. Genomic SEM Simulation Results for DHS Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

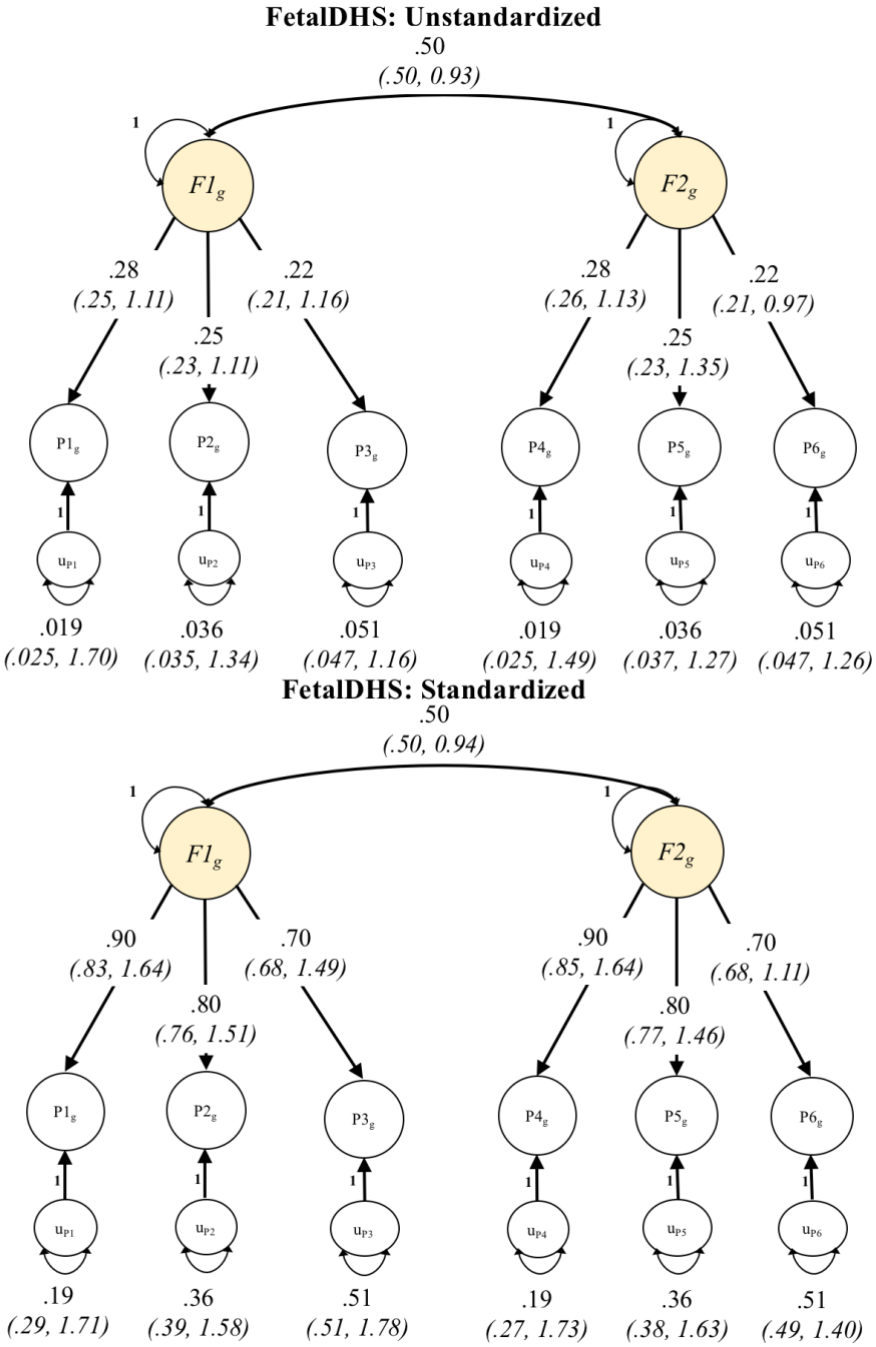

Figure 7b. Genomic SEM Simulation Results for FetalDHS Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

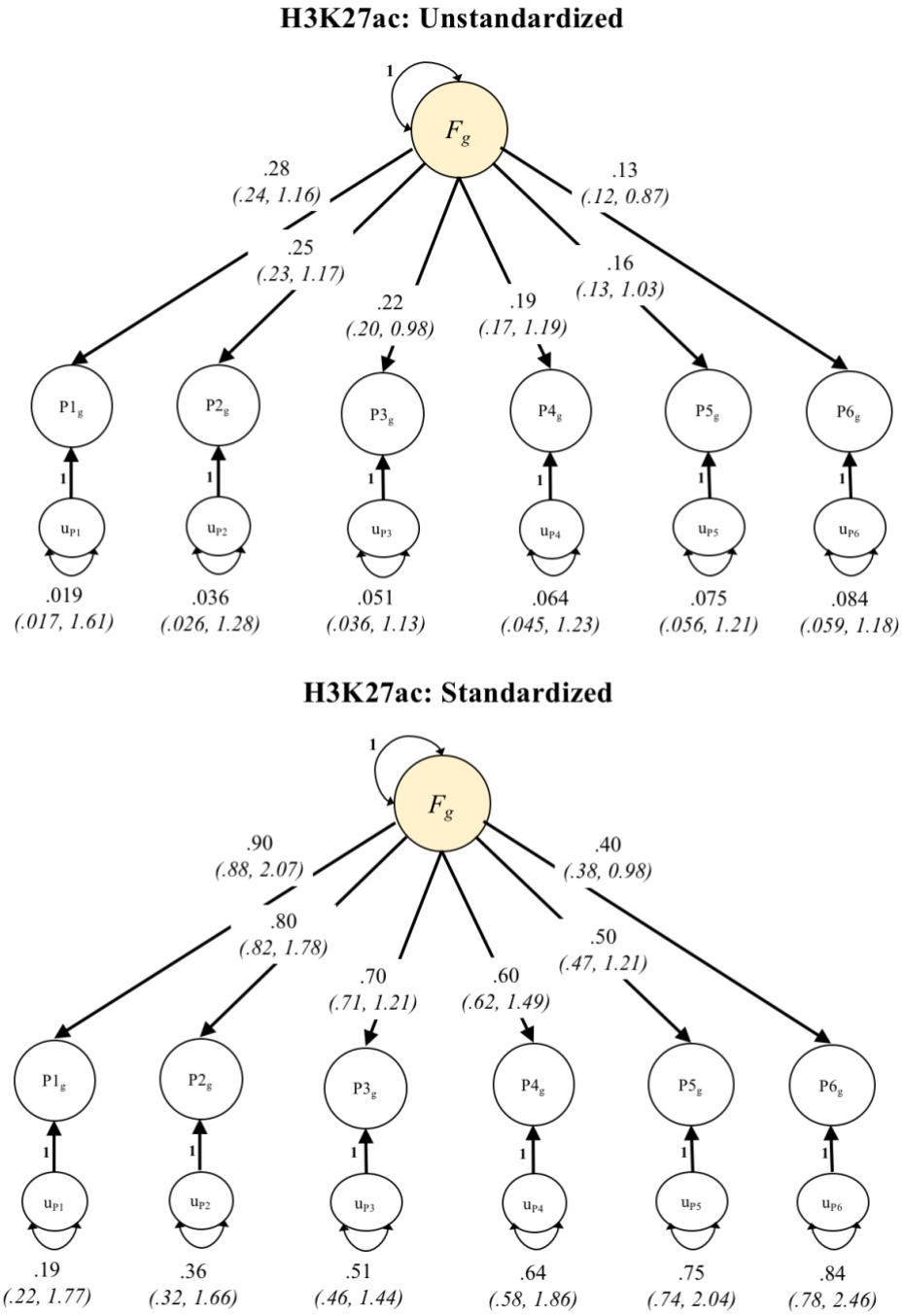

Figure 7c. Genomic SEM Simulation Results for H3K27ac Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

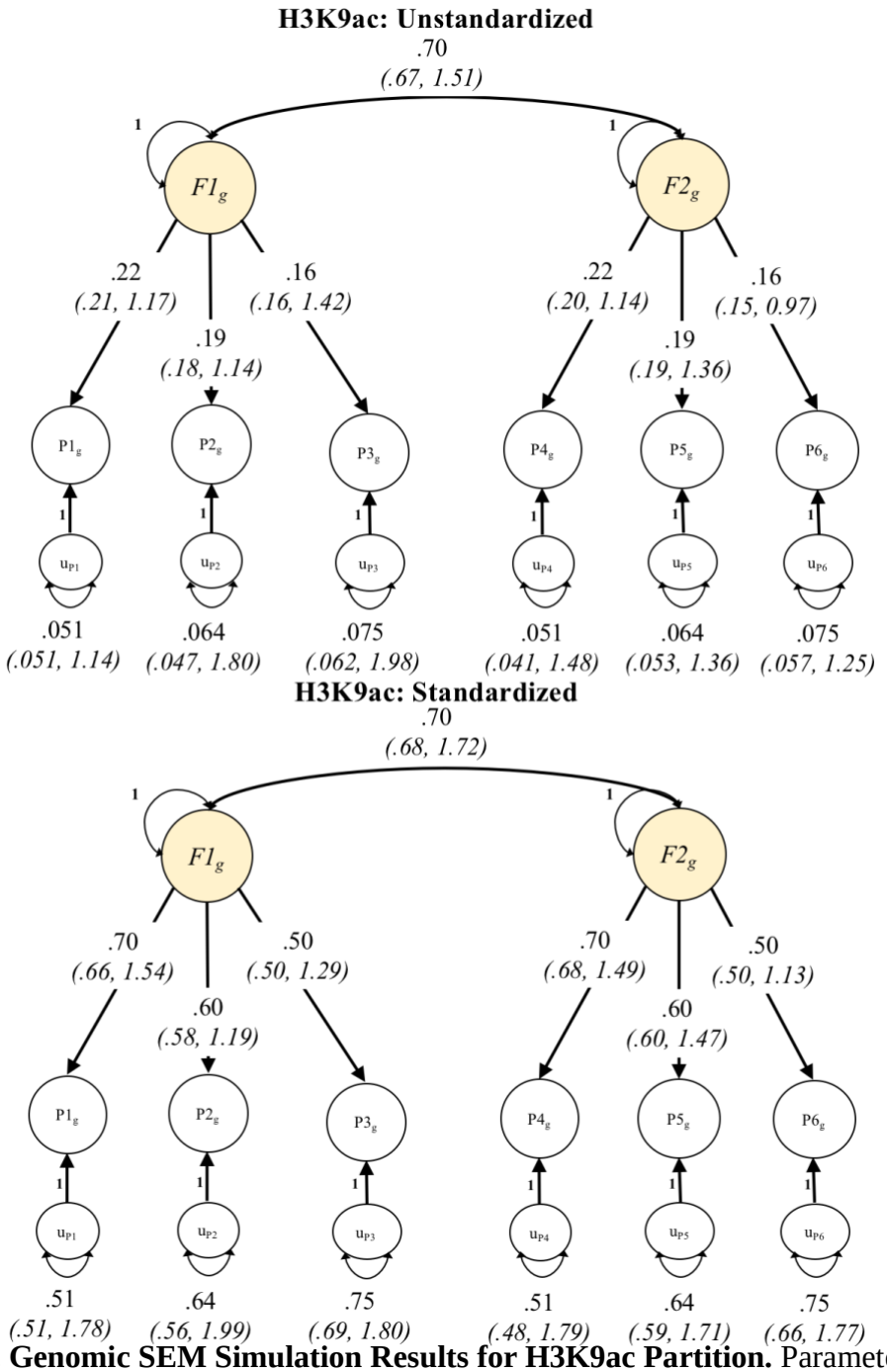

Figure 7d. Genomic SEM Simulation Results for H3K9ac Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

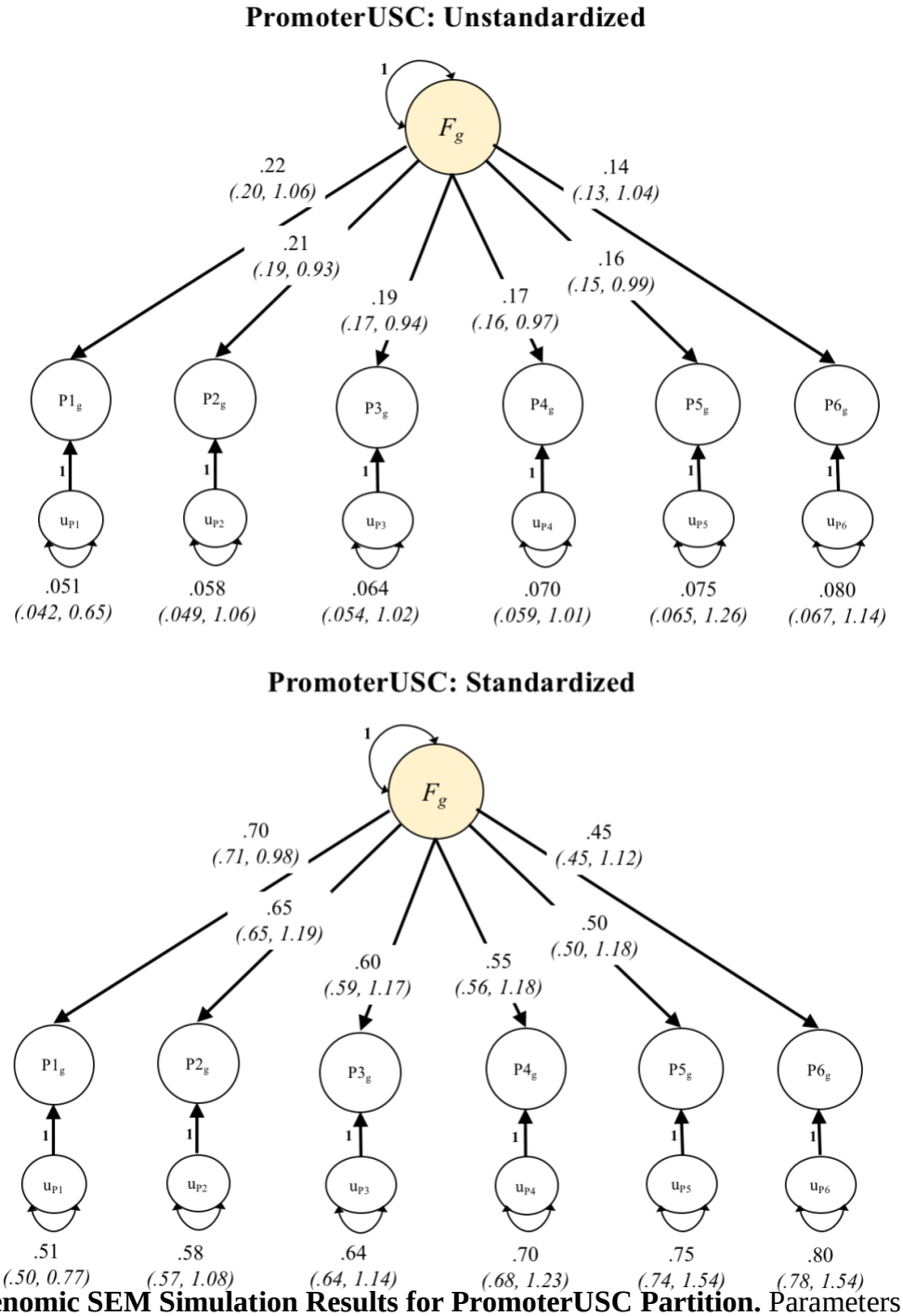

Figure 7e. Genomic SEM Simulation Results for PromoterUSC Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

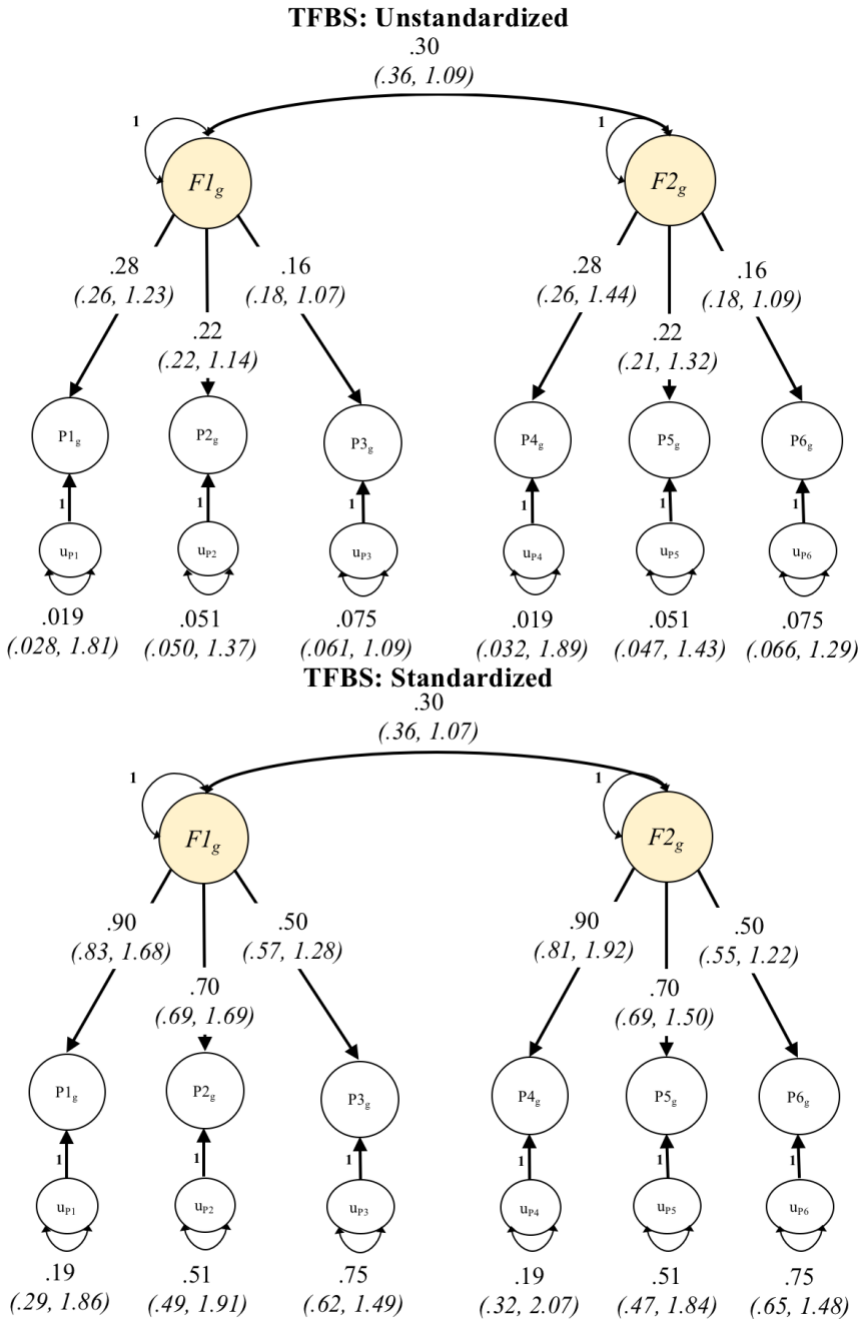

Figure 7f. Genomic SEM Simulation Results for TFBS Partition. Parameters outside of the parentheses indicate those provided in the generating population. In parentheses, we provide n the average point estimate and the ratio of the mean *SE estimate* across the 100 runs over the empirical *SE* (calculated as the standard deviation of the parameter estimates across the 100 runs). We note that *SE* estimates are expected to be upwardly biased in the standardized case, and for residual variances, due to upper or lower limits on the sampling distributions (e.g., residual variance > 0).

Table 2. Percentage of Runs Favoring Generating Model

Partition	*p*(Chi-square)	AIC	CFI
DHS Peaks	81%	89%	87%
FetalDHS	96%	82%	92%
H3K27ac	63%	72%	72%
H3K9ac	56%	56%	56%
Promoter	85%	91%	85%
TFBS	93%	89%	90%

Note. Results reflect the proportion of runs favoring the population generating model relative to a close alternative model (either a common factor or correlated factors model). For example, a common factor model was specified in the generating population for DHS peaks, and then a common and correlated factors model were compared with respect to model indices for each of the 100 runs of the simulation.

DISCUSSION

In order to utilize a Genomic SEM framework to model partitioned genetic covariance across psychiatric traits, we first validated a multivariate extension of S-LDSC. The estimation of partition-specific genetic covariance represents a substantive addition to the existing S-LDSC method, particularly with respect to understanding genetic overlap across traits. For example, it may be that the heritability of both schizophrenia and bipolar disorder is highly enriched in fetal brain cells, but that this enrichment is uninformative with respect to genetic overlap. That is, the genetic covariance within this partition may still be near 0, despite enrichment for each disorder within the partition, if the disorders are affected by a small number of disparate loci with large effects. Conversely, a partition for which heritability is less enriched for both disorders may still

be characterized by high genetic covariance if both disorders are affected by large sets of overlapping loci with relatively small effects.

After validating the multivariate extension of S-LDSC, we verified that stratified genetic covariance matrices could be as input into the general Genomic SEM framework. These results revealed that Genomic SEM produced model parameter estimates that matched the generating population, that *SEs* are estimated with minimal bias, and that model fit statistics appropriately favor the generating model within a particular partition. Pairing stratified genetic covariance matrices with a Genomic SEM framework opens up the door for a number of analytic possibilities. This includes the ability to formally model genetic enrichment a the level of overarching genetic factors. In the next Chapter, this particular application is used to identify classes of genes that are enriched for psychiatric risk factors.

Chapter 3: Application of Genomic SEM to Interrogate Genetic Overlap Across 11 Psychiatric Traits[3]

INTRODUCTION

Comorbidity is the rule rather than the exception for mental health disorders (Kessler et al., 2005). In addition to comorbidity among diagnostic categories, continuous dimensions of liability to psychopathology overlap across domains, with elevated symptoms in one continuous domain associated with elevated symptoms in other domains (Smoller et al., 2018). Recent findings indicate that this comorbidity is captured by a latent, general psychopathology factor that is commonly known as the *p*-factor and is widely supported based on previous results (Caspi & Moffitt, 2018; Caspi et al., 2013; Pettersson, Larsson, & Lichtenstein, 2016; Smoller et al., 2018; Stochl et al., 2015). Observed patterns of comorbidity might arise, in part, due to genetic variants with downstream effects on multiple mental health disorders.

The argument could be made that phenotypic comorbidity is merely a reflection of disorders causing other disorders, or key symptoms causing multiple symptom clusters (i.e., bridge symptoms in a network model), as opposed to shared biological causes. However, bivariate genomic methods such as genomic-relatedness-based restricted maximum-likelihood (GREML; Yang et al., 2011) and LDSC (Bulik-Sullivan et al., 2015) allow for the estimation of genetic correlation of mental health outcomes in largely non-overlapping participant samples. As

[3] Results reported in the third chapter appear in the preprint from (Grotzinger et al., 2020). My contributions to this work included QC and gathering of the datasets, simulations, analyses, and write-up of the results. Grotzinger, A. D., Mallard, T. T., Akingbuwa, W. A., Ip, H. F., Adams, M. J., Lewis, C. M., ... Tucker-Drob, E.M., & Nivard, M.G. (2020). Genetic Architecture of 11 Major Psychiatric Disorders at Biobehavioral, Functional Genomic, and Molecular Genetic Levels of Analysis. *medRxiv*

these methods indicate high levels of genetic overlap across mental health traits, this provides compelling evidence for common genetic correlates of psychiatric risk. However, existing molecular genetic methods do not permit the data generating processes of the observed, genetic correlations to be investigated systematically. Having developed and validate both Genomic SEM, and its extension Stratified Genomic SEM, the present chapter sought to apply these multivariate genomic methods to gain greater insight into genetic risk sharing across a wide array of psychiatric traits.

METHOD

Psychiatric Genomic Data

GWAS summary statistics were utilized for eleven major psychiatric disorders: attention-deficit/hyperactivity disorder (ADHD; Demontis et al., 2019), alcohol use disorder (AUD; Walters et al., 2018), anorexia nervosa (AN; Watson et al., 2019), autism spectrum disorder (ASD; Grove et al., 2019), anxiety disorders (ANX; Otowa et al., 2016; Purves et al., 2019), bipolar disorder (BIP; Stahl et al., 2019), major depressive disorder (MDD; Howard et al., 2018; Wray et al., 2018), obsessive compulsive disorder (OCD; International Obsessive Compulsive Disorder Foundation et al., 2018), post-traumatic stress disorder (PTSD; Duncan et al., 2017; Meier et al., 2018), schizophrenia (SCZ; [forthcoming]), and Tourette's syndrome (TS; Yu et al., 2019). Summary statistics were obtained from a range of sources, including the Psychiatric Genetics Consortium (PGC), UK Biobank (UKB), 23andMe, and iPSYCH (Table 3).

Table 3. Contributing Cohorts for 11 Psychiatric Traits

Trait	Data Source	Cases	Controls
ADHD	23andMe	5,017	57,363
	PGC	19,099	34,194
	Total	24,116	91,557
SCZ	PGC	53,386	77,258
ANX	ANGST	5,540	11,770
	UKB	25,453	58,113
	Total	30,993	69,883
ASD	PGC	18,382	27,969
BIP	PGC	20,352	31,358
MDD	PGC + 23andMe	135,458	344,901
	UKB	113,769	208,811
	Total	249,227	553,712
PTSD	iPSYCH	9,831	19,225
	PGC	23,212	151,447
	Total	33,043	170,672
TS	PGC	4,819	9,488
ALCH	PGC	8,485	20,272
	UKB	147,267	
	Total	155,697	
OCD	PGC	2,688	7,037
AN	PGC	16,992	55,525

Note. Sample sizes reported are for the European only subsets of the summary statistics. The broad depression phenotype was specifically used from the UKB Major Depression GWAS. The lifetime anxiety disorder phenotype was used for the UKB Anxiety GWAS. We note that the total sample sizes listed are the raw totals and do not reflect effective sample sizes.

The original articles contain details for the corresponding univariate GWAS about sample ascertainment, quality control, and related procedures. For five phenotypes—MDD, ADHD, ANX, PTSD, and ALCH— GWAS summary statistics were meta-analyzed across two different contributing cohorts. Meta-analyses of the same phenotype for PTSD, MDD, ADHD, ANX, and ALCH were conducted in Genomic SEM by specifying a model in which each SNP predicted a latent genetic factor defined by the two phenotypes with fixed loadings of 1. This model is statistically similar to a fixed effects meta-analysis.

Selection and Creation of Functional Annotations

In order to construct the genome-wide S-LDSC matrix, and estimate stratified genetic covariance, pre-computed annotation files provided by the original S-LDSC authors were used (Finucane et al., 2015). In line with recommendations, all annotations were included from the most recent 1000 Genomes Phase 3 BaselineLD Version 2.2 (Hujoel, Gazal, Hormozdiari, van de Geijn, & Price, 2019) that includes a total of 97 annotations ranging from coding, UTR, promoter, and flanking window annotations. For tissue specific histone marks, we included annotations constructed based on data from the Roadmap Epigenetics Project (Kundaje et al., 2015) for narrowly defined peaks for DNase hypersensitivity, H3K27ac, H3K4me1, H3K4me3, H3K9ac, and H3K36me3 chromatin. For tissue specific gene expression, annotations were included that were previously constructed based on RNA sequencing data from human tissues from GTEx (Consortium, 2015) and for annotations constructed from human, mouse, and rat microarray experiments from the Franke Lab (i.e., DEPICT; Pers et al., 2015). For both tissue specific histone/chromatin marks and gene expression, we utilized only brain and endocrine relevant regions in addition to 5 randomly selected control regions (e.g., epidermal tissue expression) from each (i.e., 10 controls total).

Twenty-nine annotations were additionally created to examine the interaction between protein-truncating variant (PTV)-intolerant (PI) genes and human brain cells. Protein-truncating variant (PTV)-intolerant (PI) genes were obtained from the Genome Aggregation Database (gnomAD), and ascertained using the probability of loss-of-function intolerance (pLI) metric. We selected genes with pLI > 0.9, producing a list of 3063 genes (Karczewski et al., 2019). Human brain cell gene sets were based on single-nucleus RNA-seq (snRNA-seq) data generated

on the Genotype-Tissue Expression (GTEx) project brain tissues (Habib et al., 2017). We selected sporadic genes and genes with low expression for the 14 cell types in the top 1600 (~15%) differentially expressed genes in each cell type, which likely cover all genes that are important for a specific cell type. PI x human brain cell gene sets contained the intersection of genes that are PTV-intolerant and each human brain cell gene set. Annotations were created using a 100kb window and LD information from the European subsample of 1000 Genomes Phase 3.

All continuous annotations and flanking window annotations were excluded for enrichment of psychiatric factors for a total of 168 binary annotations across the baseline model, gene expression, histone marks, and PI and neuronal gene annotations. For a Bonferroni correction < .05 this corresponds to $p < 2.98E\text{-}4$. The continuous annotations were retained for construction of the genome-wide, S-LDSC matrix.

Estimating Genetic Enrichment of Model Parameters

Genetic enrichment within a functional annotation reflects the extent to which the per-SNP heritability within that category is greater relative to the per-SNP heritability outside of that category. Put another way, genetic enrichment indexes whether the ratio of the proportion of heritability accounted for by a certain class of SNPs over the proportion of SNPs within that category is significantly different than the null of 1. This can be formalized as testing the null hypothesis that:

$$\text{H0: } \frac{\left(\frac{\theta|C|}{\theta}\right)}{\left(\frac{M|C|}{M}\right)} = 1$$

where $\theta|C|$ reflects the estimate for a model parameter within a functional category, θ reflects the genome-wide parameter estimate, $M|C|$ is the number of SNPs within the same functional category, and *M* is the total number of SNPs used to estimate genome-wide heritability. In order to obtain a *SE* for this ratio, enrichment is estimated for factor variances in the context of Stratified Genomic SEM using a two-step procedure. In Step 1, a model is estimated using the genome-wide S-LDSC matrix in which unit variance identification is used (i.e., factor variances fixed to 1). In Step 2, the freely estimated parameters from the previous model are fixed and the previously fixed parameter, and its *SE*, are freely estimated across the stratified genetic covariance matrices. This estimate in Step 2 reflects the proportion of the parameter of interest (e.g., psychiatric factor variance) accounted for by the genetic covariance structure within a given functional category (i.e., the numerator of the equation). This proportional estimate, and its *SE*, are subsequently divided by the proportion of SNPs in that category (i.e., the denominator of the equation). For clarification, we note that genome-wide enrichment across all SNPs is exactly equal to 1. That is, for Step 2, if the genome-wide matrix is used as input this reproduces a parameter estimate of 1, that is then divided by a proportion of 1, which reflects the ratio of *M/M* (i.e., all SNPs over all SNPs).

Quality Control Procedures

LD-Score Regression. Quality control (QC) procedures for producing the *S* and V_S matrix followed the defaults in LDSC. This included removing SNPs with an MAF < 1%, information scores (i.e., imputation quality scores) < .9, SNPs from the MHC region due to unique LD structure within this genomic area, and filtering SNPs to HapMap3 as these tend to be well-

imputed in the absence of imputation information. The LD scores used for the analyses presented were estimated from 1000 Genomes Phase 3, but restricted to HapMap3 SNPs.

Multivariate GWAS. Summary statistics are only restricted to HapMap3 SNPs for the estimation of the genetic covariance and sampling covariance matrix in LDSC, whereas all SNPs passing QC filters are included for multivariate GWAS. To obtain summary statistics for multivariate GWAS, we recommend using QC procedures of removing SNPs with an MAF < .01 in the reference panel, and those SNPs with an INFO score < 0.6. MAFs were obtained for the current analyses using the 1000 Genomes Phase 3 reference panel. Using these QC steps, 4,775,763 SNPs were present across all 11 psychiatric traits. The regression effects for the univariate indicators were standardized using the procedure for logistic coefficients outlined above.

Identification of Top Hits (Clumping) and Overlapping Hits

Lead SNPs for meta-analyzed univariate indicators and the latent genetic factors were identified using the clumping and pruning algorithm in FUMA (Watanabe, Taskesen, Van Bochoven, & Posthuma, 2017). Independent significant SNPs were defined as crossing the genome-wide significance threshold of $p < 5e\text{-}8$ that were independent from other SNPs at $r^2 < 0.1$. We used pre-calculated LD from European 1000 Genomes Phase 3 reference panel to identify independent SNPs. Top loci were subsequently identified by merging any SNPs in close proximity (< 250 kb) into a single genomic locus such that an individual locus could include multiple independent SNPs at $r^2 < 0.1$. This same pipeline was used for the full set of univariate summary statistics (i.e., not listwise deleted) in order to determine whether the factor hits were novel relative to the univariate indicators. We depict only the significant loci (referred to as hits

throughout the paper) in the manhattan plots. To identify overlap with other traits we identified all independent SNPs that were in LD ($r^2 < 0.1$) with independent SNPs for the individual traits. As LD structure can vary across different cohorts, we also considered SNPs to be overlapping/in LD with hits from the univariate studies if they were within a 250 kb window (125 kb on either side of the index variant) for loci identified for the psychiatric factors, hierarchical factor, and omnibus test described in the multivariate GWAS section below.

RESULTS

Factor Analysis of Genetic Covariance across 11 Psychiatric Traits

A heatmap of the genetic correlations across the 11 psychiatric traits clearly indicates both pervasive overlap across traits *and* tighter clustering among certain subgroups of traits (Figure 8a). Using a combination of exploratory and confirmatory models, we sought to identify the genome-wide, multivariate architecture that most closely approximated the observed patterns of genetic overlap. In order to explore the full-scope of factors solutions, EFAs were conducted using the *factanal* R package for two to five factor solutions using both oblique rotations, which allow for correlations among the latent factors, and orthogonal rotations, which assumes factors are independent (i.e., uncorrelated). Orthogonal rotations were examined in order to identify maximally separable dimensions with distinct sets of psychiatric indicators. The exploratory factor analyses (EFAs) were conducted for the genetic correlation structure derived from odd chromosomes only, excluding chromosome 23.

Confirmatory factor analyses (CFAs) specified on the basis of these EFAs were fit to a genetic correlation matrix estimated using only even chromosomes. Using odd and even chromosome covariance matrices for the exploratory and confirmatory models, respectively,

provided a form of cross-validation to guard against model overfitting. Confirmatory models based on orthogonal EFA results allowed for freely correlated factors, as pruning factor loadings will often reintroduce factor correlations. A single common factor model was also considered. All CFAs were fit using the Weighted Least Squares (WLS) estimator in the *GenomicSEM* R package.

For the CFAs, factors were assigned to traits when their standardized loading exceeded .35 in the corresponding EFAs, with two exceptions. First, for all EFAs with > 3 factors, a factor was identified with Tourette's Syndrome (TS) as its only indicator with standardized loading >.35. Assigning TS to all factors at once, or to one factor at a time, resulted in issues with model convergence. Consequently, this final factor was removed in the CFA and TS was specified to always load on the factor with the largest EFA loading (excluding the factor defined only by TS) and models were compared where TS loaded onto one of the remaining factors. Among these combinations of TS models, a final model was selected using model fit indices (i.e., AIC, SRMR, and CFI). Second, for certain EFA solutions, there were traits that did not meet the standardized loading criteria of .35 for any factor. For these traits, factors were assigned when their standardized loading exceeded a more lenient threshold of 0.2. Model fit indices were then inspected for the follow-up CFA model to confirm that including those factor loadings provided better fit to the data.

CFAs specified based on the 5-factor correlated and uncorrelated rotations were similar in both factor structure and fit to the data. In addition, both 5-factor models provided far superior fit to the data, relative to the other models, with a number of the other CFAs failing to converge (Table 4). The final model was chosen as the CFA based on the five-factor orthogonal EFA as

this ultimately provided the best fit to the data. Importantly, the model identified using a split of even and odd chromosomes also fit the data well when applied to the genome-wide matrix estimated using chromosomes 1-22 (Figure 8b; $\chi^2[33] = 161.66$, AIC = 227.66, CFI = .975, SRMR = .072).

Table 4. Confirmatory Model Comparisons

CFA based on:	*df*	AIC	CFI	SRMR
Common Factor	*44*	645.48	0.730	0.136
Oblique Rotation				
2 Factor EFA		*No Convergence*		
3 Factor EFA		*No Convergence*		
4 Factor EFA	*39*	575.23	0.767	0.109
5 Factor EFA	*37*	226.35	0.936	0.084
Orthogonal Rotation				
2 Factor EFA		*No Convergence*		
3 Factor EFA		*No Convergence*		
4 Factor EFA	*36*	309.54	0.897	0.120
5 Factor EFA	*33*	192.85	0.955	0.078
Hierarchical	*35*	235.44	0.933	0.091

Note. CFAs, for which model fit estimates are presented here, were estimated using LDSC-estimated genetic correlations for even numbered chromosomes only. They were fit based on results from EFAs for estimated using genetic correlations for odd chromosomes, excluding chromosome 23. The hierarchical model was fit to the solution based on the 5 factor orthogonal EFA.

For this model, Factor 1 largely consists of disorders characterized by compulsive behaviors (AN, OCD, TS), Factor 2 is characterized by psychotic disorders (SCZ, BIP), Factor 3 is defined primarily by childhood onset neurodevelopmental disorders (ADHD, AUT), and Factor 4 is a fundamentally internalizing factor (ANX, MDD). The fact that AUD loads on the last three factors is consistent with the psychosocially, multidetermined nature of alcohol risk. For ease, we refer to Factors 1-4 in these terms (i.e., psychotic disorders for Factor 2) in the

remainder of the text. These results largely replicate PGC Cross-Disorder Group 2 (PGC-CDG2; Lee et al., 2019) analyses based on a subset of variables, and earlier GWAS releases, of those used here. More specifically, PGC-CDG2 reported factors representing compulsive, psychotic, and neurodevelopmental disorders, which correspond closely to our first three factors. The moderate factor correlations were also suggestive of a hierarchical factor structure. Indeed, a hierarchical model also fit the data well (Figure 1c; $\chi^2[35] = 171.37$, AIC = 233.37, CFI = .974, SRMR = .079), tentatively suggesting the presence of a higher-order, general psychopathology factor (i.e., *p*-factor).

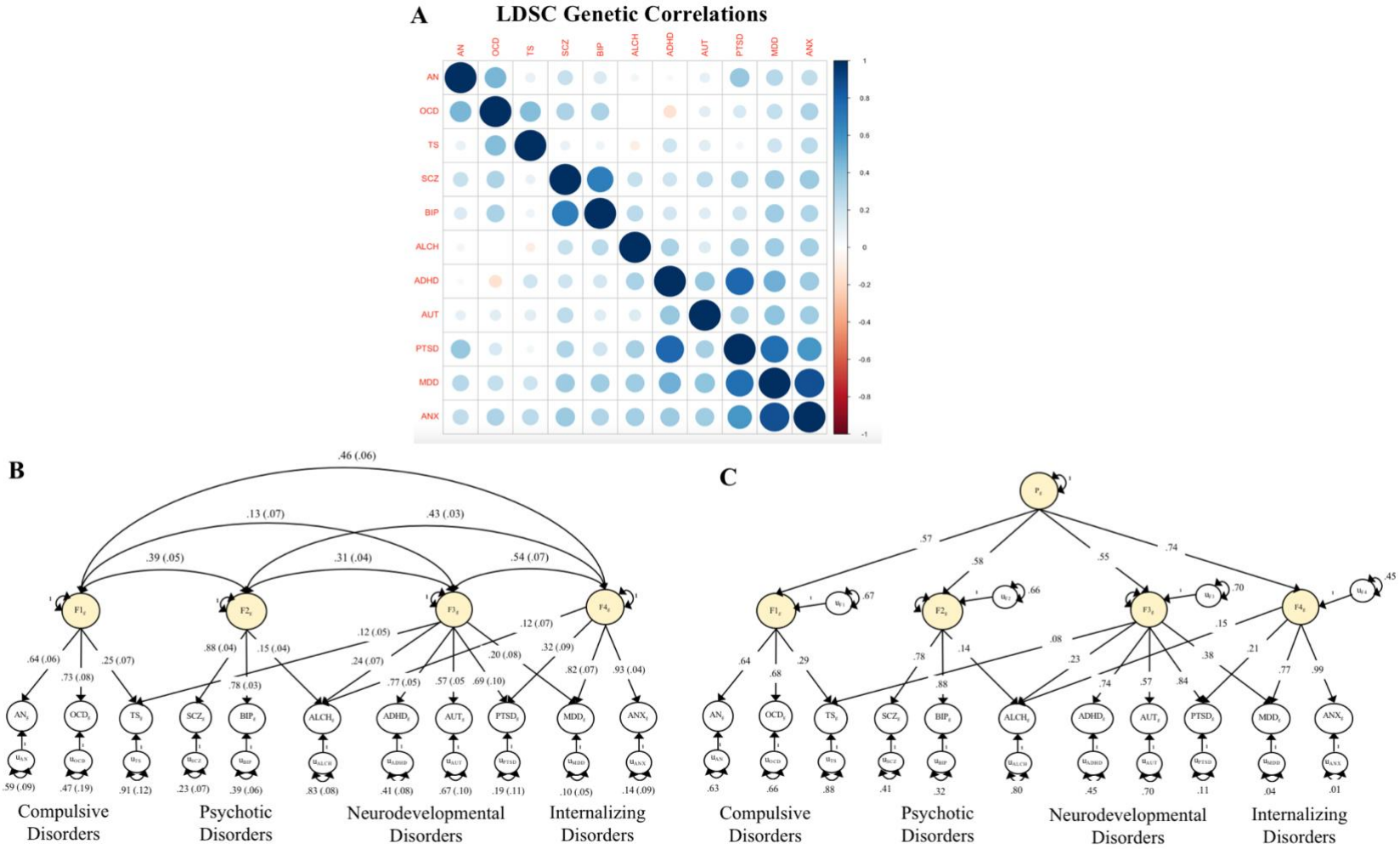

Figure 8. Heatmap of Genetic Correlations and Factor Models. Panel A: Genetic correlations estimated using LDSC. **Panel B:** Figure presents standardized results for the correlated factors model fit to the genome-wide LDSC genetic covariance matrix. **Panel C:** Figure presents standardized results for hierarchical factor model fit to genome-wide LDSC genetic covariance matrix. Standard errors are not shown for the hierarchical model as this is not currently available for models with factors as outcomes. For both Panels B and C, indicators are presented as circle to reflect the fact that these are unobserved heritability estimates. ADHD = attention-deficit/hyperactivity disorder; OCD = obsessive-compulsive disorder; TS = Tourette's syndrome; PTSD = post-traumatic stress disorder; AN = anorexia nervosa; AUT = autism spectrum disorder; ALCH = alcohol use disorder; ANX = anxiety; MDD = major depressive disorder; BIP = bipolar disorder; SCZ = schizophrenia.

Genetic Enrichment of Psychiatric Factors (Stratified Genomic SEM)

Enrichment results for the compulsive disorders factor are not considered as these findings were nonsignificant due to the relatively reduced power of the contributing psychiatric indicators (AN, OCD, TS). Overall, 40 annotations were significant using a Bonferroni correction < 5% for 168 annotations for psychotic disorders, 1 annotation (the conserved primate annotation) for neurodevelopmental disorders, 4 annotations for internalizing disorders, and 38 annotations for the *p*-factor. In line with prior findings for complex traits more generally (Finucane et al., 2015), conserved functional annotations were highly enriched across the psychiatric factors. However, the conserved primate region was particularly enriched for the neurodevelopmental, internalizing, and *p*-factor, and the Genomic Evolutionary Rate Profiling (GERP) annotation (another conserved annotation) for the neurodevelopmental and internalizing factors (Figure 9A). Also in line with findings for complex traits (Gazal et al., 2017), genetic enrichment increased slightly across the allele frequency spectrum for the minor allele frequency (MAF) bin annotations (Figure 9B). This may seem initially counterintuitive under a model of negative selection, but it is important to remember that less common variants with the same per-allele effect size as more common variants will explain *less* heritability (variance) in the population due to their low frequency (Gazal et al., 2017). While larger enrichment for lower frequency variants is possible under very strong negative selection, the present findings line up with expectations for reasonable models of selection pressure (Loh et al., 2015).

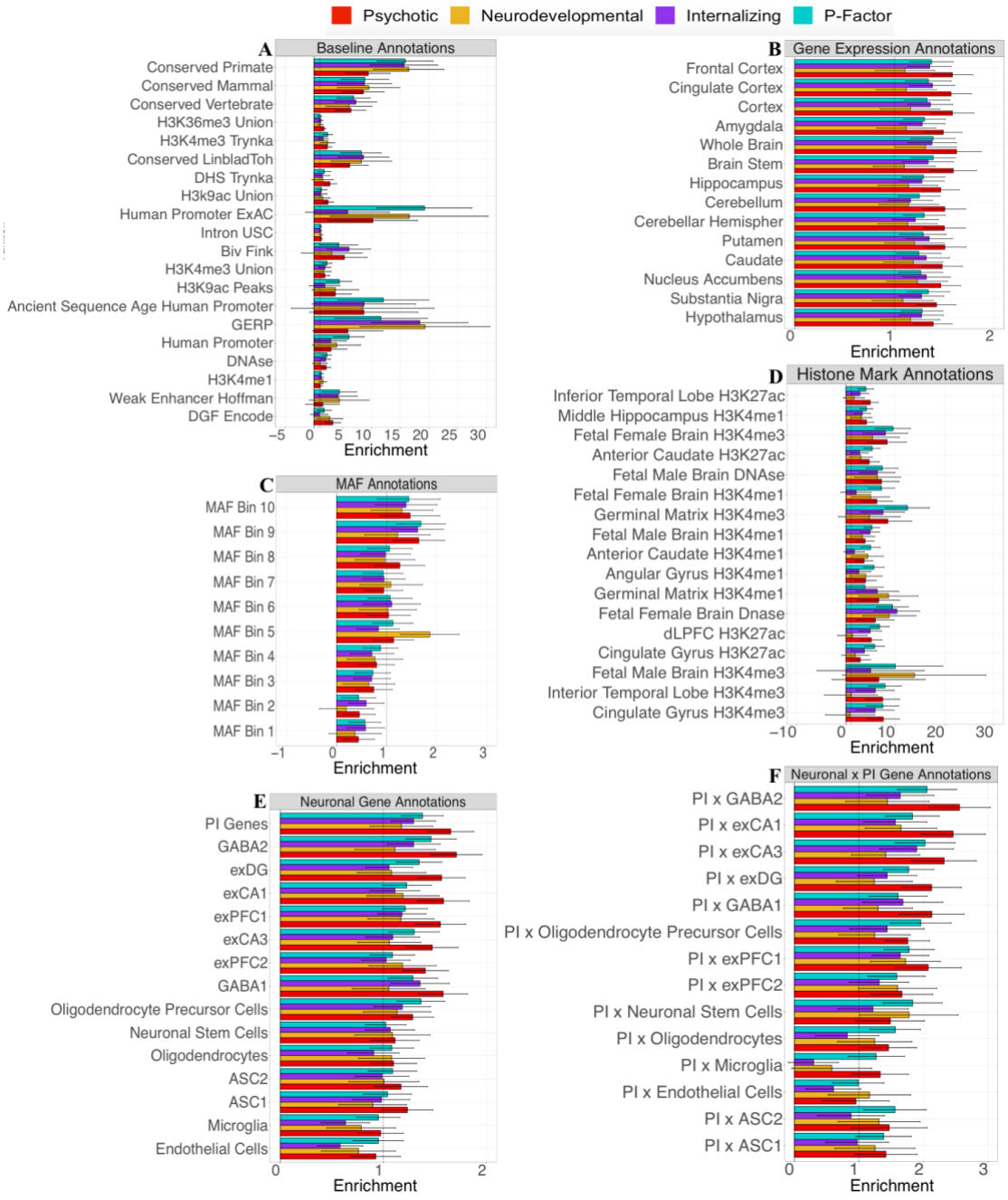

Figure 9. Genetic Enrichment of Factors. Figure depicts enrichment of psychotic (shown in red), neurodevelopmental (shown in gold), and internalizing factors (shown in purple) from the correlated factors model and the second-order p-factor (shown in light blue) for the baseline annotations (panel A), gene expression annotations (panel B), minor allele frequency annotations (MAF; panel C), histone mark annotations (panel D), neuronal gene annotations (panel E), and neuronal × PI gene annotations (panel F). Results for the compulsive disorders factor are not depicted due to the relatively low power of these results. For panels A, B, and D, only the top ten annotations across the factors are depicted within each of the functional categories. The black dashed line reflects the null (enrichment = 1). Error bars depict 95% confidence intervals.

The most significant histone marker effect for psychotic disorders was for the inferior temporal lobe H3K27ac, a region previously shown to be enriched for common epilepsies (Marzi et al., 2018). The most enriched annotation for the neurodevelopmental disorders factor was fetal female brain DNase, a region that was also enriched across the psychotic and internalizing disorders. For internalizing disorders, the most significant annotation was fetal male brain H3K4me1, a region also enriched for psychotic disorders and the *p*-factor. Both fetal female brain DNAse and fetal male brain H3K4me1 have been previously reported to be enriched for general liability across psychiatric disorders (Schork et al., 2019). As the opposite-sex counterpart for each of these annotations (e.g., fetal male brain DNAse) were also generally enriched, though to a slightly lesser degree, it is currently unclear whether this enrichment pattern offers insight into sex-specific base rates for different disorders. In line with previous results for earlier releases of schizophrenia summary statistics (Finucane et al., 2015), fetal female brain H3K4me3 region was enriched for psychotic, internalizing disorders, and the *p*-factor. The most significant histone annotation for the *p*-factor was dlPFC H3K27ac.

For gene tissue expression, we observe that brain regions are generally enriched, as is also observed for other complex traits, but were most enriched for psychotic disorders (Figure 4d; Finucane et al., 2018). As these functional categories become more refined, the enrichment signal may begin to delineate between more nuanced, neuroanatomical sub-regions and categories. To this end, 29 new annotations were created in order to examine the effects of neuronal genes, protein-truncating variant (PTV)-intolerant (PI) genes that are particularly likely to modify gene function, and their intersection. These results revealed that these functional annotations were particularly enriched for psychotic disorders, with 5 out of the 10 most

significantly enriched gene sets falling in this category. Moreover, specific intersections of PI-PTV and neuronal genes were more enriched than others, with the interaction of PI-PTV genes and GABA2 and exCA1 displaying the most significant enrichment for psychotic disorders. As PI-PTV genes have previously been found to be enriched for AUT, ADHD, BIP and SCZ (Ganna et al., 2018), these results shed light on specific neuronal subcategories that delineate between classes of disorders within PI-PTV gene sets. The final set of analyses sought to identify risk conferring loci for the psychiatric factors.

Q_{SNP} Estimation

In the context of the multivariate GWAS for the correlated factors model, four separate follow-up models were estimated in which the SNP predicted three of the overarching factors *and* the indicators of the remaining fourth factor. Comparing the model χ^2 between the model in which the SNP predicted all four factors to one of these four follow-up models produces a factor-specific $Q_{SNP.}$ Q_{SNP} indexes violation of the null hypothesis that the SNP acts through a given factor. Put another way, it quantifies whether the individual SNP is more likely to operate through the common pathways of the psychiatric factor, or the independent pathways of individual disorders. For the hierarchical factor structure, we compared the model χ^2 for a model in which the SNP predicted only the second-order, *p*-factor, to a model in which the SNP predicted only the four, first-order psychiatric factors. For the hierarchical model, Q_{SNP} indexes heterogeneity at the level of the psychiatric factors (i.e., deviation from the null that the SNP operates strictly through the *p*-factor).

Multivariate GWAS in Genomic SEM

As the schizophrenia summary statistics are not publicly available at the time of writing this dissertation, specific hits are not reported, but will be presented when submitted for publication. For the purpose of this dissertation, general trends are reported in tables and figures. To begin with, a maximally complex model in which the target SNP was allowed to have direct regression relations with each of the 11 psychiatric phenotypes was compared against a null model in which the 11 regression relations were fixed to zero. This omnibus test is χ^2 distributed with 11 *df,* and quantifies whether there is an overall effect of the SNP across the phenotypes, irrespective of the directionality of the effect. For the omnibus meta-analysis, we identified 184 hits, 39 of which were not in LD with any of the univariate hits (Table 5; Figure 10 for Manahattan plots; Figure 11 for QQ-plots). Of these 39 novel hits, 9 have not been described for outside studies of psychiatric traits/symptoms and were largely characterized by hits previously found for cognitive (e.g., intelligence) or anthropometric traits (e.g., BMI). Moreover, 7 hits were entirely novel in that they were not in LD with any previously discovered hits in the GWAS catalogue. For comparative purposes, we consider overlap with results from PGC-CDG2 (Lee et al., 2019) given both overlapping datasets and research questions. For the omnibus meta-analysis, we recapture 97 of the 146 (66.4%) PGC-CDG2 hits.

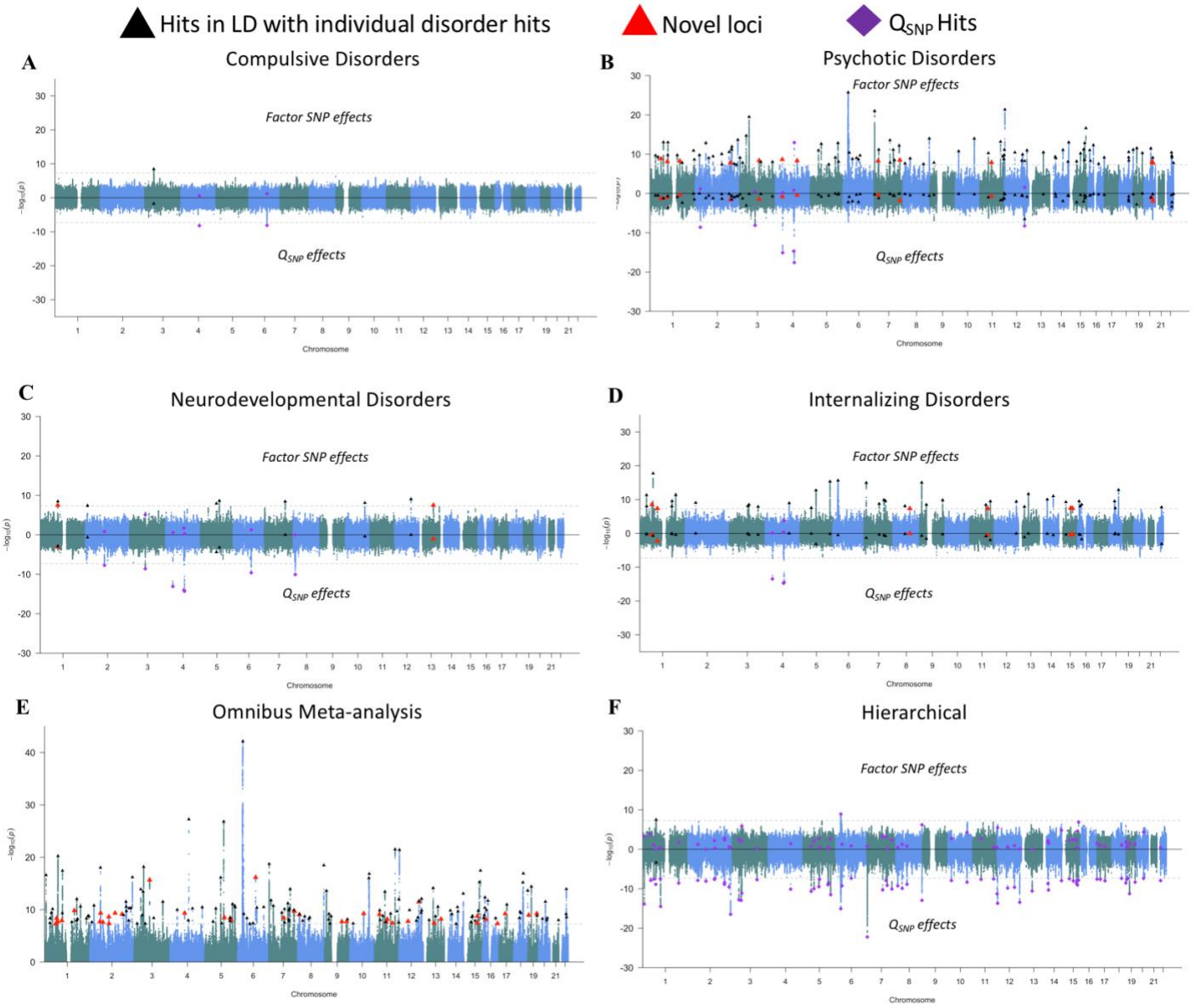

Figure 10. Manhattan Plots for Psychiatric Factors. Genomic SEM was used to conduct a multivariate GWAS from the correlated factors model for compulsive disorders (Factor 1; panel A), psychotic disorders (Factor 2; panel B), neurodevelopmental disorders (Factor 3; panel C), and internalizing disorders (Factor 4; panel D) using the genome-wide LDSC matrix. Panel E depicts results from the omnibus test across all 11 psychiatric traits. Panel F depicts the results of the SNP effect on the second-order general liability factor from the hierarchical model. The top half of the plots depicts the -log10(p) values for SNP effects on the factor; the bottom half depicts the log(10)p values for the factor specific Q_{SNP} effects. The gray dashed line marks the threshold for genome-wide significance ($p < 5 \times 10^{-8}$). Black triangles denote independent factor hits that were in LD with hits for one of the univariate indicators and were not in LD factor-specific Q_{SNP} hits. Large red triangles denote novel loci that were not in LD with any of the univariate GWAS or factor-specific Q_{SNP} hits. Purple diamonds denote Q_{SNP} hits.

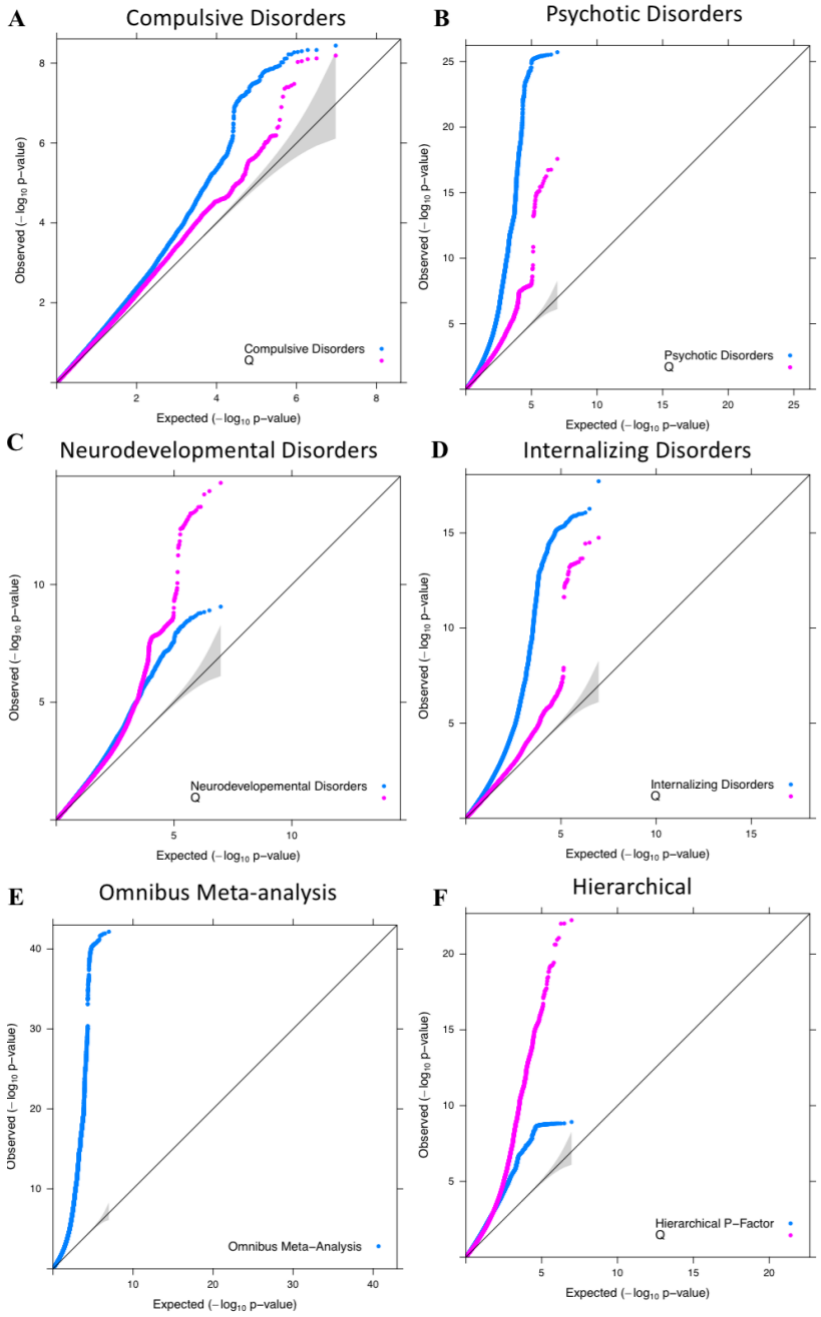

Figure 11. QQ-plots for Multivariate GWAS. Expected −log10(*p*)-values are those expected under the null hypothesis. The shaded area indicates the 95% confidence interval under the null. In the top four panels, results are shown for the compulsive disorders (panel A), psychotic disorders (panel B), neurodevelopmental disorders (panel C), and internalizing disorders (panel D) factors from the correlated factors model. Panel E depicts the results from the 11 *df* omnibus meta-analysis across all 11 psychiatric indicators. Panel F depicts results for the hierarchical, second-order factor. Blue lines depict results for the psychiatric factors. Pink lines depict the factor specific Q_{SNP} estimates.

Table 5. Genome-wide Multivariate GWAS Results

Trait	Effective *N*	Mean $\chi 2$	LDSC Univariate Intercept	Independent Hits (LD with Q hits)	LD with Trait Hits (LD with Q hits)	Unique from Trait Hits (LD with Q hits)
Multivariate GWAS						
Factor 1 (Compulsive)	19,108	1.209	0.973	1 (0)	1 (0)	0 (0)
Factor 2 (Psychotic)	87,138	1.869	0.975	108 (1)	96 (1)	12 (0)
Factor 3 (Neurodevelopmental)	55,932	1.301	1.022	9 (0)	7 (0)	2 (0)
Factor 4 (Internalizing)	455,340	1.635	0.997	44 (0)	38 (0)	6 (0)
Hierarchical	667,343	1.795	0.955	2 (1)	1 (0)	1 (1)
Meta-Analysis	-	2.216	0.883	184	145 (-)	39 (-)
Heterogeneity Index (Q_{SNP})						
Factor 1 Q_{SNP}	-	1.113	1.001	2	1	1
Factor 2 Q_{SNP}	-	1.251	0.994	6	4	2
Factor 3 Q_{SNP}	-	1.246	0.980	7	4	3
Factor 4 Q_{SNP}	-	1.142	0.977	3	3	0
Hierarchical Q_{SNP}	-	1.667	0.928	69	58	11

Trait	Effective *N* (Total *N*)	Mean $\chi 2$	LDSC Univariate Intercept	Independent Hits (LD with Q hits)	LD with Factor Hits (LD with Q hits)	Unique from Factor Hits (LD with Q hits)
AN	34,467 (72,517)	1.297	1.020	8 (0)	1 (0)	7 (0)
OCD	5,712 (9,725)	1.062	0.993	0 (0)	0 (0)	0 (0)
TS	9,614 (14,307)	1.123	1.014	1 (0)	0 (0)	1 (0)
SCZ	87,462 (130,644)	2.118	1.077	179 (2)	89 (2)	90 (0)
BIP	35,967 (51,710)	1.396	1.020	16 (0)	9 (0)	7 (0)
ALCH	155,697 (155,697)	1.199	0.994	6 (3)	2 (1)	4 (2)
ADHD	46,586 (115,673)	1.221	0.969	6 (0)	3 (0)	3 (0)
AUT	33,719 (46,351)	1.198	1.008	3 (1)	0 (0)	3 (1)
PTSD	22,001 (38,593)	1.119	0.991	0 (0)	0 (0)	0 (0)
MDD	498,520 (802,939)	1.957	1.024	109 (0)	43 (0)	66 (0)
ANX	30,273 (100,876)	1.194	0.998	2 (0)	2 (0)	0 (0)

Note. Independent hits were defined using a pruning window of 250Kb and $r^2 < 0.1$. Hits are considered in LD if their LD was $R^2>.10$ or within a 250Kb window of one another. Values in parentheses indicate whether any of the hits were in LD with hits for factor-specific Q_{SNP} hits from the respective model. Effective sample size (*N*) was estimated using the procedure outlined in the online supplement of Mallard et al. (2019).

From the correlated factors model, we identified 1 hit for the compulsive disorders factor that was also a hit for AN, and two hits for factor specific $Q_{SNP.}$ We identified 108 hits for the psychotic disorders factor, 96 of which were in LD with hits from BIP and SCZ, and 12 that were novel relative to the univariate traits. Of these 12 unique hits, 8 have been reported as hits in outside GWAS of psychiatric traits, 2 were novel for psychiatric traits, and 2 were entirely novel. Factor specific Q_{SNP} revealed 6 hits, 3 of which were in LD with hits for ALCH, including a locus in the well-described ADH1B gene that was identified as a Q_{SNP} hit for all four factors.

We identified 9 hits for the neurodevelopmental disorders factor, 3 of which were in LD with hits for ADHD or MDD, and 2 of which were novel relative to the univariate traits. These two novel hits were in LD with hits previously described for outside GWAS of psychiatric traits. There were 7 hits for the neurodevelopmental Q_{SNP} estimate, one of which was specific to AUT (rs10099100). For internalizing disorders, we identified 44 independent hits, 6 of which were not in LD with any univariate hits. Among these 6 novel loci, 3 were identified in outside studies of psychiatric traits, one has been identified for smoking initiation, and two have yet to be described for any trait. Three loci were identified for internalizing $Q_{SNP,}$ all three of which were in LD with hits for ALCH. We note that the loss of univariate MDD hits relative to the factor hits reflects a combination of signal specific to MDD and splitting the MDD signal across two factors. Of the 146 hits from PGC-CDG2, none were in LD with hits for the compulsive disorders factors, 73 hits were in LD for psychotic disorders, 5 hits for neurodevelopmental disorders, and 19 hits for internalizing disorders. As 6 of these overlapping hits were redundant across the factors, the correlated factors model recaptures 91 of the 146 (62.4%) PGC-CDG2 hits.

Hierarchical results revealed only 2 hits on the second-order *p*-factor, both of which were in LD with univariate hits for MDD and SCZ, and have been described in multiple outside studies of psychiatric traits. The *p*-factor was characterized by the highest level of heterogeneity by far, with 69 loci identified for hierarchical Q_{SNP}, 49 of which were in LD with hits on the four psychiatric factors from the correlated factors model. Using the ratio of mean χ^2 for the factor summary statistics over mean χ^2 for factor-specific Q_{SNP} estimates we produce an overarching sense of homogenous relative to heterogeneous signal at the level of the factor indicators. This ratio lent the most support to the psychotic disorders (ratio: 1.494) and internalizing disorders factor (ratio: 1.432), tentative support for the compulsive disorders factor (ratio: 1.086), and the least support for the neurodevelopmental disorders factor (1.044) and hierarchical factor (1.076). For the *p*-factor in particular, the combination of a high heterogeneity ratio and the lack of hits suggests that the ability to model a transdiagnostic risk factor reflects the convergence of intermediary processes that lie somewhere between overarching, genome-wide covariance and SNP level effects.

DISCUSSION

High levels of genetic overlap and polygenicity across psychiatric traits necessitate multivariate genomic methods that can be used to test competing hypotheses about what gave rise to the data generating process. Genomic SEM and Stratified Genomic SEM are applied here to identify an overall factor structure, genetic enrichment at the level of the psychiatric factors, and genetic loci that are likely culprits for causes of both phenotypic divergence and convergence. Analyses identify two competing factor structures. One model consisted of 4 correlated latent factors that segregate across compulsive, psychotic, neurodevelopmental and

internalizing disorders. A second model included a second-order factor that influences these four correlated factors. Consistent with the phenotypic literature, this model provided tentative evidence for a genomic *p*-factor, indexing transdiagnostic risk across psychiatric disorders. However, genome-wide estimates of genetic overlap offer limited biological insight, and may not hold for certain classes of genes or individual SNPs.

In follow-up Stratified Genomic SEM models, genetic enrichment was identified for the psychiatric factors across 168 binary, functional annotations. These results provided tentative support for both the correlated and hierarchical factor model. In support of the hierarchical model, there were some enrichment patterns that held across all psychiatric traits, including enrichment of conserved regions and brain regions more generally. Enrichment within conserved regions (e.g., conserved primate, GERP) is generally indicative of negative selection pressure for covariance patterns across psychiatric traits, whereby genetic variance with negative consequences on evolutionary fitness in these regions is selected against. Given both the general nature of evolutionarily conserved annotations, and the finding that enrichment in these is also observed across a range of complex traits (e.g., cognitive function, anthropometric traits; Finucane et al., 2018; Gazal et al., 2017), enrichment in these regions likely marks disruption in processes important for general functioning. In line with the correlated factors model, specific patterns of enrichment were also observed within the four psychiatric factors, including much stronger enrichment for annotations that index the intersection of PI-PTV and neuronal genes for the psychotic disorders factor. Multivariate GWAS was then implemented to drill down even further to individual SNP effects.

Multivariate GWAS results identified novel loci not present for the univariate traits, highlighting the potential of Genomic SEM to provide new insight into biological processes that could not be gleaned from the individual traits. However, the goal of Genomic SEM is not explicitly to boost power, but rather to identify SNPs operating through hypothesized factor structures. To this end, $Q_{SNP,}$ and the ratio of mean χ^2 for the factor effects over Q_{SNP} mean χ^2, was used to evaluate whether SNPs tended to operate through the identified factor structure. In general, results lent support to pleiotropic pathways for compulsive, psychotic, and internalizing disorders, but not for the neurodevelopmental disorders or *p*-factor. These results have important theoretical implications for a rapidly expanding *p*-factor literature, suggesting that a general factor of psychopathology does not hold at the level of individual genetic variants.

It is important to note that while Genomic SEM appropriately accounts for sample overlap across univariate GWAS in estimating statistical , it is not able to correct for case cross-contamination, in which individuals with one disorder are inappropriately diagnosed with another disorder. For example, it could be that estimates of genetic overlap between bipolar disorder and schizophrenia are upwardly biased as individuals with schizophrenia are incorrectly included in the bipolar disorder GWAS.

General Discussion

The ability to model genetic covariance matrices estimated from GWAS summary statistics produced from samples with varying and unknown degrees of sample overlap represents a substantive addition to genomic methods. Even for those not interested in genetics, hypotheses about relationships between variables that are quite rare, or exclusionary of one another (e.g., early and late onset of the same disorder), can now be examined using publicly available genomic data that can be readily downloaded online. For example, research groups with hypotheses about the relationship between early and late onset Alzheimer's could now run a multiple regression-type model that controls for the overlap in genetic signal between these two traits, while also examining unique relationships with outside traits (e.g., metabolic, cognitive, socioeconomic traits). To the best of our knowledge, such a model would not have been possible prior to the advent of a method like Genomic SEM.

In addition to examining systems of relationships at the level of genome-wide genetic overlap, these models can be extended to specific categories of genes using Stratified Genomic SEM. The ability of this method to advance etiological understanding will rapidly increase as the collateral information used to construct these categories become increasingly specific. For example, until recently, only tissue- but not cell-specific gene-expression data needed to create partitions were available. While useful for guiding future work, tissue-specific expression panels are non-specific, as they contain an array of distinct cell types and fail to provide a certain level of reductive precision. With the advent of single-cell RNA-sequencing (scRNA-seq), combined with recent surveys of cells across the central nervous system (CNS; Skene et al., 2018; Ziesel et al., 2018), even finer-grained partitions can be constructed. In an age where many psychiatric

medications are prescribed without diagnostic specificity (e.g., SSRIs; Vaswani, Linda, & Ramesh, 2003), this methodological framework has the potential to identify novel intervention targets for particularly comorbid presentations.

Genomic SEM and Stratified Genomic SEM can be utilized to examine an array of theoretical questions, but these methods were originally developed with the express purpose of providing greater understanding of extensive psychiatric comorbidity. To this end, these novel methods were used to interrogate genetic overlap across 11 psychiatric traits. As many of these disorders are quite rare, it is highly unlikely that a sufficiently sized phenotypic sample exists that includes comorbidity patterns across such a wide diagnostic spectrum. Indeed, as certain disorders are exclusionary of one another, it would actually be impossible to examine their phenotypic associations. Thus, the multivariate, genomic framework described here offers the first chance to interrogate the full spectrum of disorders described in the current diagnostic system.

At the genome-wide level, Genomic SEM revealed four, correlated psychiatric factors that are best described as reflecting compulsive, psychotic, neurodevelopmental, and internalizing disorders. These results suggest some possible avenues of revision for future iterations of the DSM (APA, 2013). For example, bipolar disorder and schizophrenia are both highly heritable disorders that cluster together quite strongly, can display similar symptom presentations (e.g., psychotic features for schizophrenia; manic episodes with psychotic features for bipolar disorder), and, therefore, may best be considered within a single disorder class (e.g., psychotic and related disorders). The argument is not that genetics should be the single determinant of how future diagnostic manuals are constructed. Rather, that genetics offers a

unique opportunity to examine the multivariate architecture of a wide array of disorders, and as such should be considered as relevant evidence for future diagnostic revisions.

Using Stratified Genomic SEM to estimate enrichment at the level of psychiatric factors, we find both patterns of divergence and convergence across disorders. Across disorders, conserved and brain regions were generally enriched for the psychiatric factors. However, the intersection of PI-PTV and neuronal genes were particularly enriched for psychotic disorders, and there was a general trend for brain regions to be slightly more enriched for psychotic disorders. This lends some tentative support to consistent findings that schizophrenia is often found to be more consistently associated with brain-based abnormalities (Shenton, Dickey, Frumin, & McCarley, 2001). As the univariate GWAS become better powered, and the categories used to construct these functional categories become increasingly specific, the methods outlined here will begin to provide an even more nuanced understanding of shared and specific risk across disorders.

At the most fine-grained level, multivariate GWAS was used to examine SNP-level effects on the psychiatric factors. Results reveal a collection of loci that were both specific to individual disorders (e.g., SNPs within the ADH1B gene for alcohol use disorders), and loci that appear to operate at the level of the psychiatric factors. This included hits that were novel relative to the univariate GWAS or outside studies of psychiatric traits. Interestingly, we find only two loci that are significant for a *p*-factor that captures genetic risk across the 11 included disorders. These results indicate that, at the very least, some of the strongest genetic signals are not operating through this general risk factor. One possibility is that the *p*-factor described in the current literature largely reflects a pattern of disorders causing disorders, or symptoms causing

symptoms, when it is phenotypically modeled for individuals within the same sample. However, to the extent that a disorder like schizophrenia typically causes life stress that subsequently leads to later onset of a disorder like depression, the depression GWAS would still pick up on some genetic signal that is shared with schizophrenia. That is, univariate GWAS signals might still pick up on effects of disorders causing disorders, yet we still find very little evidence for SNPs operating through the *p*-factor. A more reasonable explanation for this pattern of findings is that a genetic *p*-factor reflects genetic overlap for secondary, transdiagnostic risk factors that are less central to psychiatric etiology relative to more primary risk factors that segregate across the four, correlated factors identified in the present analyses. For example, it may be that dysregulation at the intersection of PI-PTV genes and certain neuronal genes (e.g., GABA2) is a primary risk factor specific to psychotic disorders, dysregulation in human promoter regions confers primary risk for neurodevelopmental disorders, and tuning of conserved regions is a secondary risk factor across disorders and complex traits more generally.

The extent to which factor-specific and transdiagnostic genetic enrichment identified in the current analyses map onto measurable phenotypic risk factors (e.g., emotion dysregulation, sensation seeking, etc.) will be an important next step for future research. In summary, the current dissertation includes the development and validation of multivariate genomic methods that are broadly applicable to an array of human complex traits. In their application to psychiatric traits, a comprehensive picture begins to emerge of patterns of genetic divergence and convergence across the diagnostic spectrum at the genome-wide level, for classes of genes, and for individual loci.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Barban, N., Jansen, R., de Vlaming, R., Vaez, A., Mandemakers, J. J., Tropf, F. C., et al. (2016). Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature Genetics*, *48*(12), 1462.

Baselmans, B. M., Jansen, R., Ip, H. F., van Dongen, J., Abdellaoui, A., van de Weijer, M. P., et al. (2017). Multivariate Genome-wide and integrated transcriptome and epigenome-wide analyses of the Well-being spectrum. *bioRxiv*, 115915.

Beaumont, R. N., Warrington, N. M., Cavadino, A., Tyrrell, J., Nodzenski, M., Horikoshi, M., et al. (2018). Genome-wide association study of offspring birth weight in 86,577 women identifies five novel loci and highlights maternal genetic effects that are independent of fetal genetics. *Human Molecular Genetics*.

Bentler, P. M., & Hu, L. T. (1995). Evaluating model fit. *Structural Equation Modeling: Concepts, Issues, and Applications*, 76–99.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., et al. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*(3), 291-298.

Bush, W. S., Oetjens, M. T., & Crawford, D. C. (2016). Unravelling the human genome–phenome relationship using phenome-wide association studies. *Nature Reviews Genetics*, *17*(3), 129.

Caspi, A., & Moffitt, T. E. (2018). All for One and One for All: Mental Disorders in One Dimension. *American Journal of Psychiatry*, appi.ajp.2018.1. http://doi.org/10.1176/appi.ajp.2018.17121383

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., et al. (2013). The p Factor. *Clinical Psychological Science*, *2*(2), 119–137. http://doi.org/10.1177/2167702613497473

Cheung, M. W.-L. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, 1521.

Consortium, G. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, *348*(6235), 648–660.

Crossley, N. A., Mechelli, A., Fusar-Poli, P., Broome, M. R., Matthiasson, P., Johns, L. C., ... & McGuire, P. K. (2009). Superior temporal lobe dysfunction and frontotemporal dysconnectivity in subjects at risk of psychosis and in first-episode psychosis. *Human brain mapping*, *30*(12), 4129-4137.

de Vlaming, R., Johannesson, M., Magnusson, P. K., Ikram, M. A., & Visscher, P. M. (2017). Equivalence of LD-score regression and individual-level-data methods. *bioRxiv*, 211821.

Dean, K., Green, M. J., Laurens, K. R., Kariuki, M., Tzoumakis, S., Sprague, T., et al. (2018). The impact of parental mental illness across the full diagnostic spectrum on externalising and internalising vulnerabilities in young offspring. *Psychological Medicine*, *48*(13), 2257–2263.

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, *51*(1), 63.

Duncan, L. E., Ratanatharathorn, A., Aiello, A. E., Almli, L. M., Amstadter, AB, Ashley-Koch, A. E., et al. (2017). Largest GWAS of PTSD (N= 20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular …*.

Elia, J., Glessner, J. T., Wang, K., Takahashi, N., Shtir, C. J., Hadley, D., ... & Lyon, G. J. (2012). Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder. *Nature genetics*, *44*(1), 78-89.

Fatemi, S. H., Reutiman, T. J., Folsom, T. D., Rooney, R. J., Patel, D. H., & Thuras, P. D. (2010). mRNA and protein levels for GABA A α4, α5, β1 and GABA B R1 receptors are altered in brains from subjects with autism. *Journal of autism and developmental disorders*, *40*(6), 743-750.

Fatemi, S. H., Folsom, T. D., Rooney, R. J., & Thuras, P. D. (2013). Expression of GABA A α2-, β1-and ε-receptors are altered significantly in the lateral cerebellum of subjects with schizophrenia, major depression and bipolar disorder. *Translational psychiatry*, *3*(9), e303-e303.

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., & Farh, K. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, *47*(11), 1228.

Finucane, H. K., Reshef, Y. A., Anttila, V., Slowikowski, K., Gusev, A., Byrnes, A., et al. (2018). Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nature Genetics*, *50*(4), 621.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466.

Forstner, A. J., Hofmann, A., Maaser, A., Sumer, S., Khudayberdiev, S., Mühleisen, T. W., ... & Treutlein, J. (2015). Genome-wide analysis implicates microRNAs and their target genes in the development of bipolar disorder. *Translational psychiatry*, *5*(11), e678-e678.

Ganna, A., Satterstrom, F. K., Zekavat, S. M., Das, I., Kurki, M. I., Churchhouse, C., et al. (2018). Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *The American Journal of Human Genetics*, *102*(6), 1204–1211.

Gao, S. F., Klomp, A., Wu, J. L., Swaab, D. F., & Bao, A. M. (2013). Reduced GAD65/67 immunoreactivity in the hypothalamic paraventricular nucleus in depression: A postmortem study. *Journal of affective disorders*, *149*(1-3), 422-425.

Gazal, S., Finucane, H. K., Furlotte, N. A., Loh, P.-R., Palamara, P. F., Liu, X., et al. (2017). Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. *Nature Genetics*, *49*(10), 1421.

Genetics, International Obsessive Compulsive Disorder Foundation, Arnold, P. D., Askland, K. D., Barlassina, C., Bellodi, L., Bienvenu, O. J., et al. (2018). Revealing the complex genetic architecture of obsessive–compulsive disorder using meta-analysis. *Molecular …*, *23*(5), 1181.

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., et al. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513.

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431–444.

Guillozet-Bongaarts, A. L., Hyde, T. M., Dalley, R. A., Hawrylycz, M. J., Henry, A., Hof, P. R., ... & Shen, E. (2014). Altered gene expression in the dorsolateral prefrontal cortex of individuals with schizophrenia. *Molecular psychiatry*, *19*(4), 478-485.

Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nature Methods*, *14*(10), 955–958.

Hibar, D. P., Westlye, L. T., Doan, N. T., Jahanshad, N., Cheung, J. W., Ching, C. R., ... & Krämer, B. (2018). Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Molecular psychiatry*, *23*(4), 932-942.

Howard, D. M., Adams, M. J., Shirali, M., Clarke, T.-K., Marioni, R. E., Davies, G., et al. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications*, *9*(1), 1470.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, *6*(1), 1–55.

Hu, W., MacDonald, M. L., Elswick, D. E., & Sweet, R. A. (2015). The glutamate hypothesis of schizophrenia: evidence from human brain tissue studies. *Annals of the New York Academy of Sciences*, *1338*(1), 38.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods*, *11*(2), 193.

Hujoel, M. L., Gazal, S., Hormozdiari, F., van de Geijn, B., & Price, A. L. (2019). Disease heritability enrichment of regulatory elements is concentrated in elements with ancient sequence age and conserved function across species. *The American Journal of Human Genetics*, *104*(4), 611–624.

Hyman, S. E. (2018). The daunting polygenicity of mental illness: making a new map. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1742), 20170031.

Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., et al. (2018). Genome-wide Analysis of Insomnia (N= 1,331,010) Identifies Novel Loci and Functional Pathways. *bioRxiv*, 214973.

Kaplan, D. (2008). Structural equation modeling: Foundations and extensions, *10*.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., et al. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*, 531210.

Kenny, D. A. (2014). Measuring model fit.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617–627.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539), 317–330.

Lee, J. J., & Chow, C. C. (2017). LD Score regression as an estimator of confounding and genetic correlations in genome-wide association studies. http://doi.org/10.1101/234815

Lee, P. H., Anttila, V., Won, H., Feng, Y.-C. A., Rosenthal, J., Zhu, Z., et al. (2019). Genome wide meta-analysis identifies genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *bioRxiv*, 528117.

Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., Perlis, R. H., et al. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, *45*(9), 984.

Lloyd-Jones, L. R., Robinson, M. R., Yang, J., & Visscher, P. M. (2018). Transformation of summary statistics from linear mixed model association on all-or-none traits to odds ratio. *Genetics*, genetics. 300360.2017.

Loh, P.-R., Bhatia, G., Gusev, A., Finucane, H. K., Bulik-Sullivan, B. K., Pollack, S. J., et al. (2015). Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nature Genetics*, *47*(12), 1385.

Mallard, T. T., Linnér, R. K., Okbay, A., Grotzinger, A. D., de Vlaming, R., Meddens, S. F. W., et al. (2019). Not just one p: Multivariate GWAS of psychiatric disorders and their

cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. *bioRxiv*, 603134.

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., et al. (2017). A general psychopathology factor (P factor) in children: structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal …*, *126*(1), 137.

Martin, N. G., & Eaves, L. J. (1977). The genetical analysis of covariance structure. *Heredity*, *38*(1), 79.

Marzi, S. J., Leung, S. K., Ribarska, T., Hannon, E., Smith, A. R., Pishva, E., et al. (2018). A histone acetylome-wide association study of Alzheimer's disease identifies disease-associated H3K27ac differences in the entorhinal cortex. *Nature Neuroscience*, *21*(11), 1618–1627.

McLaughlin, K. A., Gadermann, A. M., Hwang, I., Sampson, N. A., Al-Hamzawi, A., Andrade, L. H., et al. (2012). Parent psychopathology and offspring mental disorders: results from the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, *200*(4), 290–299.

Meier, S., Trontti, K., Als, T. D., Laine, M., Pedersen, M. G., Bybjerg-Grauholm, J., et al. (2018). Genome-wide Association Study of Anxiety and Stress-related Disorders in the iPSYCH Cohort. *bioRxiv*, 263855.

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., et al. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular …*, *21*(10), 1391.

O'Reilly, P. F., Hoggart, C. J., Pomyen, Y., Calboli, F. C., Elliott, P., Järvelin, M.-R., & Coin, L. J. (2012). MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS ONE*, *7*(5), e34861.

Pers, T. H., Karjalainen, J. M., Chan, Y., Westra, H.-J., Wood, A. R., Yang, J., et al. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*, *6*, 5890.

Pettersson, E., Larsson, H., & Lichtenstein, P. (2016). Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. *Molecular …*, *21*(5), 717.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.

Purves, K. L., Coleman, J. R., Meier, S. M., Rayner, C., Davis, K. A., Cheesman, R., et al. (2019). A major role for common genetic variation in anxiety disorders. *Molecular …*, 1–12.

Ray, D., & Boehnke, M. (2018). Methods for meta-analysis of multiple traits using GWAS summary statistics. *Genetic Epidemiology*, *42*(2), 134–145.

Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling: a Multidisciplinary Journal*, *21*(1), 149–160.

Savalei, V., & Bentler, P. M. (2009). A two-stage approach to missing data: Theory and application to auxiliary variables. *Structural Equation Modeling: a Multidisciplinary Journal*, *16*(3), 477–497.

Schork, A. J., Won, H., Appadurai, V., Nudel, R., Gandal, M., Delaneau, O., et al. (2019). A genome-wide association study of shared risk across psychiatric disorders implicates gene regulation during fetal neurodevelopment. *Nature Neuroscience*, *22*(3), 353–361.

Schmaal, L., Hibar, D. P., Sämann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., ... & Vernooij, M. W. (2017). Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Molecular psychiatry*, *22*(6), 900-909.

Selzam, S., Coleman, J. R., Caspi, A., Moffitt, T. E., & Plomin, R. (2018). A polygenic p factor for major psychiatric disorders. *Translational Psychiatry*, *8*(1), 205.

Shenton, M. E., Dickey, C. C., Frumin, M., & McCarley, R. W. (2001). A review of MRI findings in schizophrenia. *Schizophrenia research*, *49*(1-2), 1-52.

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2018). Psychiatric genetics and the structure of psychopathology. *Molecular …*, 1.

Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., et al. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, *51*(5), 793.

Stochl, J., Khandaker, G. M., Lewis, G., Perez, J., Goodyer, I. M., Zammit, S., et al. (2015). Mood, anxiety and psychotic phenomena measure a common psychopathological factor. *Psychological Medicine*, *45*(7), 1483–1493.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. *Sage Focus Editions*, *154*, 10–10.

Turley, P., Walters, R. K., Maghzian, O., Okbay, A., Lee, J. J., Fontana, M. A., et al. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics*, *50*(2), 229.

Van der Sluis, S., Posthuma, D., & Dolan, C. V. (2013). TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genetics*, *9*(1), e1003235.

Vaswani, M., Linda, F. K., & Ramesh, S. (2003). Role of selective serotonin reuptake inhibitors in psychiatric disorders: a comprehensive review. *Progress in neuro-psychopharmacology and biological psychiatry*, *27*(1), 85-102.

Verhulst, B., Maes, H. H., & Neale, M. C. (2017). GW-SEM: A Statistical Package to Conduct Genome-Wide Structural Equation Modeling. *Behavior Genetics*, *47*(3), 345–359.

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., et al. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, *21*(12), 1656.

Watanabe, K., Taskesen, E., Van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1), 1826.

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R., Gaspar, H. A., et al. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, *51*(8), 1207–1214.

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*(5), 668.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, *88*(1), 76–82.

Yu, D., Sul, J. H., Tsetsos, F., Nawaz, M. S., Huang, A. Y., Zelaya, I., et al. (2019). Interrogating the genetic determinants of Tourette's syndrome and other tic disorders through genome-wide association studies. *American Journal of Psychiatry*, *176*(3), 217–227.

Yuan, K. H., & Bentler, P. M. (2000). Robust mean and covariance structure analysis through iteratively reweighted least squares. *Psychometrika*, *65*(1), 43–58.

Zhu, X., Feng, T., Tayo, B. O., Liang, J., Young, J. H., Franceschini, N., et al. (2015). Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *The American Journal of Human Genetics*, *96*(1), 21–36.